2018

Relationships of Gender on Mathematics Achievement in High Achieving Military-Connected Children

Randi Kay Rose Dalton
Walden University

Walden University

College of Education

This is to certify that the doctoral dissertation by

Randi Kay Rose Dalton

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Deanne Otto, Committee Chairperson, Education Faculty
Dr. Jesse Richter, Committee Member, Education Faculty
Dr. Jennifer Lapin, University Reviewer, Education Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2018

Abstract

Relationships of Gender on Mathematics Achievement in High Achieving Military-Connected Children

by

Randi Kay Rose Dalton

MA, Emporia State University, 1997

MA, Emporia State University, 1996

BS, Emporia State University, 1993

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Curriculum, Instruction, and Assessment

Walden University

November 2018

Abstract

The purpose of this quantitative correlational study was to describe the relationship between gender and grade level to mathematics achievement for high achieving military-connected students in Grades 3 through 9 who attended American public schools between 2012 and 2016. The theoretical framework was based on Sax's research on gender differences in learning. The research questions were: if there was a statistically significant difference between the percentage of military-connected students scoring in the top 2 quartiles (at or above the national average) for normal curve equivalent (NCE) mathematics scores on the Terra Nova Third Edition (TNTE) using gender as a predictor; and if there was a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles for NCE mathematics scores on the TNTE by gender and grade band. The sample size consisted of archival scores from 135,571 students, aggregated into 136 representative grades and provided by the participating school district's Research Center. A two-tailed t test was conducted to answer Research Question 1. The results were *alpha* = .05, t(-.696), df = 134, and p = .000. An ANOVA and logistical regression were conducted to answer Research Question 2, *alpha* = .05, F(.168),and p =.984. There were no statistical differences between the mean numbers of females and males by grade level or grade band. There was gender equity within the population studied. Therefore, the null hypotheses were accepted for both research questions. This study contributes to positive social change by adding to the limited body of knowledge about mathematics achievement for high achieving military-connected students relative to gender.

Relationships of Gender to Mathematics Achievement

in High Achieving Military-Connected Children

by

Randi Kay Rose Dalton

MA, Emporia State University, 1997

MA, Emporia State University, 1996

BS, Emporia State University, 1993

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Curriculum, Instruction, and Assessment

Walden University

November 2018

Dedication

I dedicate this doctoral study to my amazingly supportive, humorous, and compassionate husband, Jeremy, who took over a myriad of responsibilities without complaint and to my awesome children, Christopher, Andria, and Tabitha, who were extremely supportive and funny during this journey. I also dedicate it to all of the wonderful educators in the world who are making the world a better place daily and specifically to two ladies who have made a tremendous positive impact on me personally as well as countless others, Marilyn Dalton and Gail Meadows-Livingstone. Finally, I dedicate it to Bob and Wendy, truly the best parents and role-models in the world. They taught me to cherish education and have always been my biggest cheerleaders. Thank you for instilling in me the love of learning and the desire to always endeavor to be the best I could be and to strive make a positive impact in our global community. You taught me compassion, persistence, and joy. Thank you also to all of my family and friends around the world for your love and support. May the force be with you!

Acknowledgments

This project coincided with a tri-continental adventure which originated during my voyages throughout Asia, continued as I journeyed through North America, and finally culminated during my explorations of Europe. The dissertation itself has been an extensive and often arduous journey of a different kind and I could never have succeeded without a tremendous amount of support. First and foremost, I would like to thank Dr. Deanne Otto, my chair, for always being a much-needed guide. She was always supportive and extremely helpful. I would also like to thank the other members of my team, Dr. Jesse Richter and Dr. Jennifer Lapin, for reading this dissertation and providing valuable feedback and Dr. Htway for his very informative statistics tutoring sessions. This would not have been possible without Dr. Deanna Boddie, my former chair who unfortunately retired before I finished, but nevertheless was instrumental in helping me to see that there was a light at the end of the tunnel, and constantly assuring me it was not a train. Their scholarly advice was invaluable as I traversed this immense pathway to success.

Thank you again to my husband, children, friends, extended family, and parents for love, hugs, support, kindness, sympathy, and for not letting me give up when the dissertation writing process was frustrating. Finally, I am very appreciative of my dogs, Koko and Shadow, who spent hundreds of hours on the floor warming my feet and providing emotional support instead of running around outside chasing squirrels and deer. My sincere thanks and appreciation go to all of you.

Table of Contents

List of Tables

Chapter 1: Introduction to the Study

Approximately 1.4 million American children live in families with at least one active duty military parent while an additional 23% of the 20 million veterans have children under the age of 18 (Creech & Hadley, 2014). Approximately 1.2 million children of active duty members are of school age (De Pedro et al., 2011; Department of Defense Education Activity, 2015). Approximately 90% of the nation's military-connected students attend civilian public schools (De Pedro, Esqueda, Cederbaum, & Astor, 2014).

Assessment of high achieving military-connected student performance at the national level in American public schools has been limited. In relation to measuring student achievement in mathematics for students in the United States at the national level, the National Assessment of Educational Progress (NAEP) is the only test that has been administered to representative populations of public and private school students across all 50 states since the 1990s (National Center for Education Statistics, 2015; O'Gara & Kanelis, 2015).

Discrepancies in mathematics achievement between male and female students are of concern because mathematical understanding is important in maintaining the nation's economic competitive advantage in a global environment (Birenbaum, Tatsuoka, & Xin, 2005). The necessity for increasing educational opportunities related to careers in science, technology, engineering, and mathematics (STEM) in the United States has been well documented (Committee on STEM Education: National Science and Technology Council, 2013; Krishnamurthi, Bevan, Rinehart, & Coulon, 2013; Laursen, Thiry, Archie,

& Crane, 2013; Tyler-Wood, Ellison, Lim, & Periathiruvadi, 2012; Wilkerson & Haden, 2014). Currently, the United States Congress annually funds a variety of STEM initiatives with the goal of furthering social justice, competing in global markets, and educating citizens for STEM related occupations (Committee on STEM Education: National Science and Technology Council, 2013). However, females are still underrepresented in these STEM fields (Mosatche, Matloff-Nieves, Kekelis, & Lawner, 2013; Tyler-Wood et al., 2012).

This study was needed, therefore, because little research has been conducted about the relationship of gender to mathematics achievement for high achieving military-connected students. This study contributes to positive social change by adding to the body of knowledge about the relationship of mathematics achievement for high achieving military-connected students relative to gender. Society stands to benefit from improved knowledge of gender equity in high achieving military-connected students in the area of mathematics, in particular for female students, because student understanding of mathematical concepts is integral to attaining success in science and other related fields (Weiss, Pasley, Smith, Banilower, & Heck, 2003).

In this chapter I include the background for this study, which contains a summary of the research literature. In addition, the research problem, purpose of the study, the research question and hypotheses, and the theoretical framework for this study are described. A brief explanation of the methodology of the study, including the research design and rationale, is also presented. Additionally, terms relevant to this study, assumptions, scope and delimitations, limitations, and the significance of this study are included.

Background

Despite the fact that approximately 90% of the nation's 1.3 million students who are designated as children of military parents attend civilian public schools (De Pedro et al., 2014), there is limited data on their academic performance because educators in civilian schools do not routinely monitor their enrollment or academic achievement (Esqueda, Astor, & De Pedro., 2012). In addition, a lack of research exists about how educators in civilian schools meet the learning needs specific to military students (De Pedro et al., 2014; Esqueda et al., 2012). Educators in civilian public schools have traditionally not had appropriate support systems in place to help these students with military-specific stressors, such as parental deployment, that often negatively impact their lives (Angrist & Johnson, 2000; Bradshaw, Sudhinaraset, Mmari, & Blum, 2010; Esqueda et al., 2012). In a study about children of military service members, De Pedro et al. (2011) found that military children in public schools are often not identified or monitored. In addition, public school teachers may not be trained in understanding the military culture and its effects on military students. In a study about the children of military service members, De Pedro et al. found that over 50% of the teachers do not adequately understand military family culture, and over 90% have not received training on how to work with these children. In this study, teachers reported frustration in not understanding the military chain of command and how to use a military school liaison, if one was available, to help them navigate the military infrastructure (De Pedro et al., 2011). De Pedro et al. concluded that a cultural disconnection exists between public school systems and children of military service members and therefore more research is needed about how to improve instruction for these students. In a study about the impact of work-related absences on families during the Gulf War, Angrist and Johnson (2000) found that parental deployments and other stressors, such as frequent moves and changes in schools, are positively associated with higher levels of depression, anxiety, stress, and reduced academic functioning.

Historically, De Pedro et al. (2011) noted that research on military children has been limited to clinical or retrospective samples and has been conducted in contexts such as the Vietnam War or the conflicts in Iraq or Afghanistan. Some prior research exists on the education of children of military parents, but much of it addresses how the military lifestyles of parents affect their children's health, behavior, or socioemotional wellbeing. In a systematic review of the impact of military deployment and reintegration on parenting, Creech and Hadley (2014) found several studies that revealed correlations between parental deployment and socioemotional problems that military children develop, including self-destructive behaviors such as alcohol and drug abuse. Creech and Hadley also noted that children of parents returning from deployment whose parents were diagnosed with depression or Post Traumatic Stress Disorder (PTSD) often had emotional health problems, such as depression. Creech and Hadley recommended that more research should be conducted in relation to gender because female military children might suffer from these emotional health problems in greater numbers than male military children. Further research is warranted because even though both genders between the ages of 5 and 12 experience an increased level of anxiety during parental deployment, young female children exhibit more externalized anxiety (Creech & Hadley, 2014). These expressions of externalized and internalized stresses may affect the genders differently during the adolescent years as well. Creech and Hadley suggested that adolescent male students may experience increased emotional distress and physical abuse from caregivers, while adolescent female students may experience more emotional abuse, particularly at the beginning and end of deployments. Gender-related stress management strategies that teachers use within the classroom environment might be improved with further research in this area. In other related research, Moeller, Culler, Hamilton, Aronson, and Perkins (2015) conducted a literature review of 26 qualitative and quantitative studies of the effects of military-related parental absences on school-aged children and found that military students with deployed parents are more likely to demonstrate behavioral problems than their civilian or military peers. Similarly, children of deployed parents are more likely than their civilian or military peers to experience difficulties with academic functioning, such as lower test scores and lower grades.

A few researchers have also examined educator or parent perceptions about the schoolwide impact of military-connected student enrollment on academic and socioemotional learning. Garner, Arnold, and Nunnery (2014) examined the perspectives of staff members at public schools with high military populations and found that staff members believed that school reform is needed to adequately meet the needs of military children. Garner et al. recommended that public school staff be trained in how to meet the unique needs of military children. They also recommended coordination among professionals across schools to improve their academic success and to provide social and emotional support. In a related study about the effects of parental combat deployment on military children and at-home spouses, Lester et al. (2010) interviewed 272 children of parents who had recently returned from months of deployment and their parents. Lester et al. found that female military children with a currently deployed parent exhibited significantly higher anxiety, depression, and other emotional distress behaviors than their male peers. Lester et al. also found that children's stress levels correlated with parental stress levels and that return from deployment sometimes led to new stressors. Lester et al. concluded that parental deployment could be a predictor of anxiety, depression, and other emotional distress behaviors for female military children. Thus, male and female students are affected by the deployment of their parents in different ways, and, therefore, their engagement in classroom instruction related to STEM, and ultimately to STEM career fields, may also be affected.

Problem Statement

In recent decades, growth in the number of females entering professional careers in science and mathematics has been documented, particularly in areas related to medicine and biological sciences (Ceci & Williams, 2010, 2011; National Science Foundation, 2017; Stoet & Geary, 2018; Wang, Eccles, & Kenny, 2013). More than half of the medical doctorates and PhDs in life sciences, psychology, and veterinary medicine are earned by females (Ceci & Williams, 2010). Yet in more mathematically intensive fields, such as physics, the number of females continues to lag behind the number of males (Ceci & Williams, 2011; Cheryan, Ziegler, Montoya, & Jiang, 2017; Johnson, Barnard-Brak, Saxon, & Johnson, 2012; Jones, 2010; National Science Foundation, 2017; Stoet & Geary, 2018). At the university level, females are still earning less than 20% of degrees in mathematically intensive areas, including computer sciences and physics (Cheryan et al., 2017). Gender discrepancies in STEM fields are a global phenomenon. Contini, Tommaso, and Mendolia (2017) noted that gender differences are widespread in the 37 Organization for Economic Cooperation and Development countries, and male students tend to perform better than female students. Females continue to be underrepresented in STEM fields other than life and social sciences globally (Stoet & Geary, 2018). In the United States, the deficiency of female engagement in STEM fields that requires advanced mathematics has remained fairly constant for decades (National Science Foundation, 2017). In an analysis of an international database containing the 2015 Programme for International Student Assessment (PISA) scores in science for 519,332 students from 72 countries, Stoet and Geary (2018) found that in almost all countries that more females appeared to be adequately prepared and capable to achieve STEM degrees than chose to graduate with STEM degrees. In addition, Stoet and Geary found that countries with high gender equality had larger gender gaps in STEM college degrees than countries with less gender equality. This pattern was repeated in many countries throughout the world. For instance, Finland, Norway, and Sweden have high gender equality, high educational performance in relation to other countries worldwide, and a small to nonexistent gender gap in adolescents. However, these countries have large gender gaps favoring males in STEM college degrees.

Lack of mathematical ability is not the fundamental reason for the underrepresentation of females in STEM careers (Ceci, Williams, & Barnett, 2009; Eccles, 2009; Wang, Eccles, & Kenny, 2013). Female students routinely outscore male students in mathematical coursework (Jones, 2010). However, females are perennially in the minority in relation to technology and mathematics careers as well as science and engineering careers that are mathematically intensive (Johnson et al., 2012; Jones, 2010). Because females are capable of attaining mathematics related careers and are capable of making mathematical contributions to society, yet they are not, this problem needs to be investigated in order to address this gender imbalance in the unique military-connected student population. Identifying when the gender gap begins is an important step towards understanding and correcting this gender imbalance. Among highest achieving students, the gender gap increases as students grow older (Contini et al., 2017). Mathematical confidence for high achieving females in STEM majors is low (Sax, Kanny, Riggers-Piehl, Whang, & Paulson., 2015). The lack of confidence in mathematics ability appears in high school also despite demonstrated mathematical aptitude (Sax et al., 2015). Among high achieving students who scored in the top 10% of the SAT mathematics test, females were less likely to rate their abilities as in the top 10% no matter what major they were in, even mathematics majors (Sax et al., 2015).

Female high school students are more likely to graduate in the top 10% of their class and earn a higher grade-point average than male high school students, yet they are also more likely to earn a lower score on the mathematics portion of the Scholastic Aptitude Test (SAT) college entrance examinations (College Board, 2007, 2015).

Female students drop out of STEM courses at the university level earlier and more frequently than male students, but the reasons are not fully understood (Spitzer & Aronson, 2015). Researchers do not agree on the causative factors for this leak in the STEM pipeline (Johnson et al., 2012; Jones, 2010). Female students may drop out of STEM courses earlier and more frequently than male students due to a lack of motivation or for cultural, social, or educational reasons (Spitzer & Aronson, 2015). Other possible causative factors include a negative self-assessment of their mathematics abilities (Johnson et al., 2012), negative attitudes towards problem solving (Cutumisu & Bulut, 2017), high mathematical anxiety (Beilock, Gunderson, Ramirez, & Levine, 2010), or stereotypical beliefs that male students are better than female students in mathematics (Beilock et al., 2010; Johnson et al., 2012; Wang & Degol, 2017). Wang and Degol (2017) posited that the gender gap in STEM careers could be reduced by breaking down cultural barriers, reducing misinformation, and reducing stereotypes. Sax et al. (2015) found that lower mathematical self-concept and subsequent confidence in mathematical abilities is still a predictor of females' pursuit of STEM careers. However, over the last 4 decades, it has had less impact (become less salient), meaning it could still be part of the explanation for the underrepresentation of females in STEM fields, but this is not a complete explanation.

Implicit bias against female students in schools and in the workplace could also be a contributing factor (Beilock et al., 2010; Jones, 2010). Another possible causative factor could be lack of exposure to professional female role models, including a lack of female mathematics teachers, particularly at secondary schools and universities. For example, 94% of preprimary teachers and 87% of primary teachers in the United States are female while only 62% of secondary teachers are female (UNESCO Institute for Statistics, 2015). The female to male ratio in the general population is nearly 1:1, consisting of 49% females to 51% males (United States Census Bureau, 2012). However, the female to male ratio in the United States military is dramatically different. Female military-connected students are part of the overall larger military communities where only 17% of the total service members (active duty and selected reserve) are female compared to 83% of the total service members who are male (Military Onesource, 2015). Thus, female military-connected students may be at a particular disadvantage regarding professional female role models if they live within a military community. Similarly, teacher expectations for female students might be lower, which might lead to teachers encouraging fewer females to pursue more mathematics instruction. Martinez et al. (2016) found that preservice teachers had lower expectations for female student mathematic achievement than male students. Martinez et al. also found that the anxiety that preservice teachers felt towards mathematics influenced the preservice teachers' beliefs about student achievement and need for academic support in mathematics. Martinez et al. suggested that mathematics anxiety of preservice teachers could negatively impact their abilities to create mathematical environments to meet the needs of both genders.

Yet another possibility of the gender gap is that the highest achieving mathematics students are the most likely to choose STEM majors in college and subsequent STEM careers, and the gender gap in this group favors males (Contini et al., 2017; Lindberg, Hyde, Petersen, & Linn, 2010; Stoet & Geary, 2013). Lindberg et al. (2010) examined data from the NAEP from the previous 2 decades and found that a national gender gap favoring males at the highest achievement levels. Similarly, Stoet and Geary (2013) examined data from the PISA for approximately 1.5 million students internationally and also found a gender gap favoring males at the highest level (95^{th} percentile and above) of mathematical achievement. Therefore, this study was needed because a gap in the current research literature exists about the relationship of gender to the mathematics achievement of high achieving military-connected students.

Purpose of the Study

The purpose of this quantitative study was to describe the relationship between gender and mathematics achievement on the Terra Nova Third Edition (TNTE) for high achieving military-connected students in Grades 3 through 9 who were enrolled in schools in the participating school district from 2012 to 2016.

Research Questions and Hypotheses

A logistic regression analysis, a two-tailed *t* test, and an ANOVA were conducted to answer the following research questions and related hypotheses, which were based on the theoretical framework and the literature review for this study.

Research Question (RQ) 1: Is there a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for normal curve equivalent (NCE) mathematics scores on the TNTE using gender as a predictor?

$H_0$1: There is no statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE using gender as a predictor.

H_a1: There is a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE using gender as a predictor.

RQ2: Is there a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE by gender and grade band?

$H_0$2*:* There is no statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE by gender and grade band.

H_a2*:* There is a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE by gender and grade band.

Theoretical Framework

The theoretical framework supporting this study was a meta-analysis of research that Sax (2005, 2009) conducted about gender differences in learning. Based on a review of current research, Sax (2005) contended that preprogrammed genetic differences exist between female and male brains. These genetic differences, Sax believed, account for both functional and anatomical differences that teachers need to address in the classroom. Sax (2005) cited research that female students use both sides of their brains to process language, but male students rely on the left side of their brains to process language (Gur et al., 1999; McGlone, 1980). Sax contended that male and female students can exhibit functional differences because anatomical differences in eyesight (Ogueta, Schwartz, Yamashita, & Farber, 1999; Wickham et al., 2000) and hearing (Cone-Wesson & Ramirez, 1997) exist between genders even in young individuals. Social interests (Connellan, Baron-Cohen, Wheelwright, Batki, & Ahluwalia, 2002) and social interactions like play preferences (Alexander & Hines, 2002) are exhibited even in infants and the differences increase as children age (Corso, 1959, 1963). In addition, Sax (2005) cited research that male children do not hear quiet sounds as well as female children (Corso, 1959, 1963). Sax (2010) also cited research that male children have a higher tolerance level for loud noises than female children (D'Alessandro & Norwich, 2009; Elliott, 1971; McGuinness, 1974; Rogers, Harkrider, Burchfield, & Nabelek, 2003). These researchers have found that rods and cones in the eyes and differences in neural pathways between the cerebral cortex and retinas allow female children to perceive their environment differently than male children (Kaplan & Benardete, 2001; Meissirel, Wilker, Chalupa, & Rakic, 1997). Female children pay more attention to texture, color, and details, which is evident in their drawings. In contrast, the rods and cones of male children are wired to pay more attention to actions, directions, locations, and speed, which is also evident in their drawings.

Sax (2005) also cited research that male children often play more aggressively and engage in fighting more often than female children; however, this type of behavior often helps male children become friends with each other (Lever, 1976, 1978).

Alternatively, female children are often engaged in passive play, such as pretend babysitting (Maestripieri & Pelka, 2002). In relation to academic subjects such as mathematics, Sax (2005) cited research that specific regions of the brain activated during visual spatial navigation and spatial-cognition performance are different for male and female students (Grön, Wunderlich, Spitzer, Tomczak, & Riepe, 2000).

By school age, Sax (2005) argued that female and male students have developed different learning styles. Sax (2005) cited studies in which researchers found that female students are more likely to try to please their teachers (i.e. by doing homework) than male students (Pomerantz, Altermatt, & Saxon, 2002; Pomerantz & Saxon, 2001; Valeski & Stipek, 2001). Male students are often not interested in pleasing their teachers unless the specific topic appeals to them (Pomerantz et al., 2002; Pomerantz & Saxon, 2001; Valeski & Stipek). This theoretical framework is described in more detail in Chapter 2, including how this research is articulated in current gender research related to mathematics achievement.

Nature of the Study

This study was quantitative in nature and used a correlational design. A correlational design involves measuring two or more variables to determine if they are associated in such a way that an increase in one variable leads to an increase or decrease in the other variable, if one variable can predict another, or if the variables are not associated (McLeod, 2008). The rationale for choosing this design was related to the purpose of this quantitative study, which was to describe the relationship between gender and grade level to mathematics achievement for high achieving military-connected students in Grades 3 through 9 who attended schools in the participating school district during the years 2012 to 2016. The independent variables were gender and grade level. The dependent variable was the percentage of military-connected students in Grades 3 through 9 who were enrolled in the participating school district during the years 2012 to 2016 and scored in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE. In relation to the methodology of this study, data collected were de-identified by the research administrator of the participating school district's Assessment Data Center, which was located on the East Coast of the United States. The district requested that they not be named within this study. I collected the number of students who scored at or above the national average for NCE mathematics scores on the TNTE for students at all of the participating schools located in the district by grade level. I did not have access to individual student names or which scores came from the individual schools from which the data originated. I collected data from the spring 2012 to the spring 2016 TNTE assessments to include 5 years of data. I analyzed the percentage of students scoring at each of the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE using analysis of variance (ANOVA), a two-tailed *t* test, and a logistic regression analysis. NCE scores were chosen because they were reported at equal intervals. I analyzed the data by individual grade level and by the following two grade bands: Grades 3 through 5 and Grades 6 through 9 in order to represent elementary and post-elementary levels. The independent variables were gender and grade band, and the dependent variable was the percentage of military-connected students in Grades 3 through 9 who were enrolled in the participating school district during the years 2012 to 2016 and scored in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE.

Definition of Terms

Deployment: Deployment refers to a temporary move of an active service military member, member of the National Guard, or Reserve member from a home base or home station to another location of military importance in the United States or overseas. Service members may be deployed with their entire unit, squadron, or ship, or they may be deployed with a small group as part of an augmentation of another unit that has already deployed. The service member's dependent family remains behind throughout the deployment, which can last from a few weeks to several months or even a whole year at a time. Service members are subject to repeated deployments. Service members may also be deployed to war zones (Military.com, 2015). Deployment commences when the military member departs for an extended mission and terminates when the member returns and is reintegrated with their family (Creech & Hadley, 2014).

Gender-based instructional strategies: These strategies are intended for female students or male students independently or for only one gender within a homogeneous or a heterogeneous environment (Sax, 2005, 2009, 2011). Examples of these strategies include providing students with teachers and role models of the same gender as the students or offering single gender classrooms or single gender schools (Sax, 2009). Sax (2011) also cited research that indicates male students learn better if they have opportunities to physically manipulate objects while female students learn better if they have opportunities to ask questions and discuss their ideas with other students (Lever, 1976, 1978; Maestripieri & Pelka, 2002).

High mathematics achievement: For this study, high mathematics achievement is defined in relation to the total mean NCE scores in mathematics that students enrolled in the participating schools in Grades 3 through 9 achieved on the TNTE from 2012 to 2016 and specifically to those students who scored at or above the national mean (within the top two quartiles). TNTE norm-referenced scores are reported as mean national stanines, grade mean equivalents, mean normal curve equivalents, mean scale scores, and median national percentiles (CTB/McGraw-Hill, 2009).

Military-connected children: For this study, military-connected children are defined as children of active duty service members in the United States military branches, which include the Air Force, Air Force Reserve, Air National Guard, Army, Army Reserve, Army National Guard, Coast Guard, Coast Guard Reserve, Marine Corps, Marine Corps Reserve, Navy, and Navy Reserve (De Pedro et al., 2011; Military.com, n.d.).

Stereotype lift: Stereotype lift, or a boost in test performance, can occur in a nontargeted group when the test is stereotype-relevant and the targeted group is subject to a negative stereotype (Marx & Stapel, 2006; Walton & Cohen, 2003). This term is important to this study because female students in Grades 3 through 9 might be subject to negative stereotypes or discrimination in relation to mathematics instruction and assessment.

Stereotype threat: Stereotype threat is a state of mind describing how stigmatized individuals do poorly on standard tests when they perceive a negative stereotype exists of their race, gender, socioeconomic status, or other real or perceived grouping Such a threat may cause individuals to perform in a way that substantiates the stereotypic expectations (Johnson et al., 2012).

Terra Nova Third Edition: This standardized, norm-referenced assessment is a product of the CTB/McGraw-Hill Company. It is an achievement test designed to describe students' strengths and weaknesses in several academic areas and is available in online and paper versions. The complete battery of tests covers mathematics, reading, language, science, and social studies. The mathematics component assesses the following areas: number and number relations, computation and numerical estimation, measurement, geometry and special skills, data analysis, statistics, and probability, patterns, functions, and algebra, and problem solving and reasoning skills (CTB/McGraw-Hill, 2009). The mathematics component is administered to students in Grades 1 through 12 in paper format and to students in Grades 2 through 12 in an online format (CTB/McGraw Hill, 2009).

Assumptions

This study was based on several assumptions. The first assumption was that the data that the participating school district's data center provided was consistent and accurate, which was important to the reliability and validity of this study. The second assumption was that the majority of students in the sample population each year spent all or most of the year in the school in which they tested, which means they were taught mathematics based on that school system's mathematics standards and adopted curricular materials. This assumption was important because student test scores needed to reflect the mathematics curriculum and instruction they received in that school district's schools. The third assumption was that external factors influencing the success of military-connected students in school impacted the genders equally. Stressors such as parental deployment, proximity to areas of civil unrest, physical separation from stateside family members and friends, and other factors associated with living in overseas military communities were assumed to influence students of both genders similarly in relation to their performance in schools generally and their performance in mathematics specifically. This assumption was important because if these stressors impacted genders differently, differences in mathematics scores could be caused by those stressors.

Scope and Delimitations

The scope of this study was defined and narrowed or delimited by the population sample and by the theoretical framework. This study was first narrowed by focusing on gender and mathematics. This study was further narrowed by focusing on high achieving military-connected students who were enrolled in the participating school district's schools, which varied from year to year and grade level to grade level due to the mobility of the military. The scope of this study was also limited to mathematics results from the TNTE for students in Grades 3 through 9 who were enrolled in the participating school district's schools from 2012 to 2016. The theoretical framework also narrowed this study, which was based on Sax's (2005, 2009) research about how genetic, anatomical, and functional differences between male students and female students impacts learning.

Sax (2005, 2009) contended that male and female students demonstrate different learning styles by school age and that female students perform better in most age groups and subjects than male students.

Concerning generalizability, examination of the participating school district was a legitimate avenue for making generalizations applicable to larger homogenous systems (Yin, 2009, 2012). Even though instruction may vary, military-connected students in the participating school district's schools were subject to the same curricular standards and assessments as their peers in other locales. Therefore, the results of this study might be generalized to children of military service members in other school systems.

Limitations

This was a small correlational study examining gender equity or inequity within a military-connected student population from the participating school district within a specific 5-year time frame and if that equity or inequity changed from grade level to grade level or within grade bands. This study included several limitations. The first limitation was related to the correlational design of this quantitative study. Numerous variables could have impacted differences in mathematics achievement between military-connected children besides grade level and gender (the variables in this study) at the time of testing. These variables might have included the following: whether the student had a parent currently deployed or soon to be deployed; if a parent was in a current war zone; branch of service; if a student was in a dual-military home; education level of parents; marriage status of parents; age of parents; number of siblings; birth order; poverty levels; race/ethnicity of students and family members; amount of time students spent overseas; number of schools students had moved to; amount of time students had lived at their current stations; whether students attended schools with a high military-connected student population; if students lived on or off base or in the United States or overseas; English Language Learner status; reading levels; physical, mental, and emotional health; and Special Education needs. There may be other factors not listed that might also have impacted test scores. Because I used de-identified secondary data from a mobile student population, I had no way of knowing how many students moved into or out of the participating school district for each school year. I also did not have access to information about specific subgroups of students at each school that were determined by poverty, race or ethnicity, English as a Second Language, or special education. Data from these subgroups might have been useful in explaining variations in scores. Untested variables might also have accounted for differences in scores. In addition, errors in my hypotheses could have led to misinterpretations of the data analysis. Descriptive validity could have been a problem because I had no way to ascertain variables such as the testing environment, how many students had deployed parents or other stressors in their lives at the time they took the test, or what other environmental factors might have occurred during the administration of the test.

This study could also be limited by my potential biases as a single researcher, which could have influenced the results of this study. Researcher bias could include promoting one theory over another or neglecting to consider multiple interpretations of the data analysis. I needed to make sure that I did not purposely or unintentionally alter the data to match my personal beliefs. In addition, because I was a teacher at a school in the participating school district, I needed to minimize my biases about that component of the school system.

These limitations are addressed in Chapter 3 in the section about threats to validity. In that section, threats to the external validity of this study, such as inaccurate inferences or overgeneralizing the results are discussed. In addition, threats to the internal validity of this study, such as making inaccurate generalizations or not considering other variables that could have influenced the test scores, are discussed. A threat to internal validity included selection parameters because the students selected for this study might not mirror the results of military-connected students from schools outside the targeted study population or nonmilitary-connected students. Another potential threat to internal validity was maturation. For example, changes in test scores for different years or different grades could be due to changes in the age of the participants and not their gender. Student attrition could also be a threat to internal validity. Threats to construct validity and statistical validity are described in this section.

Significance

The significance of this study was related to advancing knowledge, improving practice, and contributing to positive social change. This study adds to the existing body of knowledge about the relationship of gender to achievement in mathematics for high achieving military-connected students in Grades 3 through 9. Illumination of trends or non-trends in gender equity specifically related to mathematics achievement in these two grade bands could assist educators of high achieving military-connected students in seeking out other research with the goal of improving teaching practices. This study also advances knowledge about overall student achievement in mathematics for American high achieving military-connected students.

In addition, this study advances knowledge about gender equity and grade levels in the field of mathematics education, specifically for high achieving military-connected students. The findings of this study might help educators in schools with significant populations of military-connected students to consider whether or not gender-based instructional strategies are appropriate and/or relevant for mathematics instruction and assessment. The findings in this study may be relevant to educators seeking to learn more about gender issues related to mathematics achievement, particularly in relation to military-connected students.

This study also contributes to positive social change by helping researchers, educators, and parents develop an understanding about gender issues related to mathematics achievement for high achieving students in schools with significant military populations. Because gender inequity was not revealed in this study, interested researchers may be motivated to design future studies to examine why the participating school district has achieved gender equity in terms of mathematics achievement so that other school systems may benefit. The results of this study provide more information to future researchers interested in gender equity at different grade levels.

Increasing the number of Americans entering STEM fields is a national priority (Obama, 2005; United States Congress, 2015). Studies such as this one contribute to the knowledge base about gender equity in mathematics instruction and achievement. Society benefits from gender equity in mathematics instruction and achievement because achieving gender equity could eliminate gender gaps in higher education STEM degrees and STEM career fields. Increasing the number of females earning STEM degrees could also lead to an increase the number of females entering the STEM pipeline, thus increasing America's competitiveness in the global economy.

Summary

This chapter was an introduction to this study. This chapter included background information regarding current research related to mathematics achievement and gender issues for military-connected students. The sample for this study was also described, which included high achieving military-connected students in Grades 3 through 9 who were enrolled in the participating school district's schools and who completed the TNTE during the years 2012 to 2016. This chapter also included a description of the problem and the purpose of this study. Research questions and hypotheses were included, and the methodology was described, including participants, data collection, and data analysis. The theoretical framework for this study was also described, which was based on research about gender differences in learning that Sax (2005, 2009, 2010) analyzed. This chapter also included definitions for specific terms, assumptions, the scope and delimitations, limitations, and significance.

In chapter 2 I included a review of the research literature related to this study and an analysis of current research about mathematics achievement and gender differences in learning in relation to military-connected students. I concluded chapter 2 with a discussion of the themes and gaps found in the review.

Chapter 2: Literature Review

Introduction

Historically, fewer women than men have engaged in STEM careers, and women have had less exposure to the educational opportunities needed to engage in STEM careers (Tolley, 2003). Internationally, the participation of women in STEM careers in highly developed countries has increased (Stoet, Bailey, Moore, & Geary, 2016). However, regardless of recent societal improvements in gender equity, women remain underrepresented in STEM careers (Stoet et al. 2016). In the United States, although more women attend colleges and universities than men, women continue to be underrepresented in STEM fields (Alon & Gelbgiser, 2011; DiPrete & Buchmann, 2013; Mann & DiPrete, 2013; Morgan, Gelbgiser, & Weeden, 2013; National Science Foundation, 2011; Reilly, Neumann, & Andrews, 2014). Gender inequity is prevalent in students' choices of STEM majors because fewer women than men choose STEM majors that require higher-level mathematics courses (Alon & Gelbgiser, 2011; Mann & DiPrete, 2013; Morgan et al., 2013). The existence of gender inequity in relation to STEM educational and career opportunities also negatively impacts the narrowing of the gender gap in earnings because high paying STEM careers in physics, electrical engineering, and computer science continue to be male dominated (Halpern et al., 2007; Spitzer & Aronson, 2015). In addition, mathematics scores on the PISA have declined for students in the United States from a ranking of 25th in 2009 to 31st in 2012 in comparison to other countries (Spitzer & Aronson, 2015). A gender gap in the United States also favors male students at the highest achievement levels of the PISA and NAEP mathematical tests

(Lindberg et al., 2010; Stoet & Geary, 2013). A significant gap in the research literature exists in relation to mathematics achievement and gender for military-connected students. Therefore, the purpose of this quantitative study was to add to the literature by describing the relationship between gender and grade level to mathematics achievement for high achieving military-connected students in Grades 3 through 9 who attended schools located in the participating school district from 2012 to 2016.

A brief summary of the research reviewed for this chapter establishes the relevance of this problem. In relation to gender differences in mathematical learning, several studies are significant. Reilly et al. (2014) examined NAEP mathematics data for 1,925,100 students in Grades 4, 8, and 12 from 1990 to 2011 and found that the gender gap for mathematics was small but favored male students for the last 2 decades. Reilly et al. also discovered that the gap at the middle school level was widening and was largest at the highest achieving category "with males being overrepresented by a factor of 2:1" (p. 645). However, in a meta-analysis of research trends in gender and mathematics performance, Lindberg et al. (2010) examined the mathematics performance of over 1 million students of American, Canadian, European, Australian/New Zealander, Asian, African, Latin American, and Middle Eastern nationalities. After noting that gender differences for elementary and middle school students were negligible, reached a peak for high school students, and then declined for college age students and adults, Lindberg et al. found no overall gender differences in mathematics performance. Lindberg et al. also found that gender-integrated classrooms did not put female students at disadvantage in relation to mathematics achievement compared to male students. In contrast, Stoet and

Geary (2013) examined PISA data for approximately 1.5 million 15-year-old students from 75 countries from 2000 to 2009 and found that the gender gap favored male students. Stoet and Geary noted that the gap was largest for students scoring at the 95th percentile and above where male students consistently outperformed female students. Similarly, Contini et al. (2017) found that the gender gap increased with age and was largest among the highest achieving students. Stoet and Geary concluded that the gender gap in STEM fields could be related to the fact that students entering STEM fields often score in the highest percentiles on mathematics tests, and, therefore, the gender gap has remained stable over the last 2 decades.

This chapter is a review of the literature. This chapter includes a description of the search strategies used to conduct this review and a comprehensive discussion of the major tenets of the theoretical framework and how they are articulated in current research. The literature review is organized according to the following topics: (a) gender differences in mathematics learning, (b) mathematics achievement in public schools in the United States, and (c) mathematics achievement in schools with high populations of military-connected students. I conclude the chapter with a summary and a discussion of major themes that emerged from the review.

Literature Search Strategies

I used several strategies to conduct this literature review. I began by conducting a manual search at a local library for books and recent peer-reviewed journals about mathematics achievement for K to 12 students and related gender differences in learning. I also conducted a systemic online search of current literature on military students enrolled in DoDEA schools in relation to mathematics achievement and gender differences in learning because DoDEA schools are unique due to their high populations of military-connected children, DoDEA staff have been specially trained to deal with the unique stressors related to military-family life, and DoDEA students routinely score better than the national average despite having large populations of students who may experience stressors such as frequent mobility and deployed parents (Abell, 2004; Beardsley, 2015; De Pedro et al., 2011; Department of Defense Education Activity, 2010, 2011a; Esqueda et al., 2012; National Center for Education Statistics, 2015). Although I did not cite any dissertations, I read several articles found in the reference sections of several dissertations. I also conducted forward searches using Google Scholar and reviewed related articles that were cited in the original articles.

At first, I limited my search of current scholarly articles to full text, peer-reviewed journals published after January 1, 2010. However, I expanded my search to articles published after January 1, 1990 because I wanted to obtain a historical perspective about research related to mathematics achievement and gender differences in learning for K to 12 students. The specific databases included EBSCO Host, Eric, ProQuest Central, SAGE, Science Direct, Thoreau, and the Walden University Dissertation Database. Key words used for these searches included *children of military personnel + education, families of military personnel, gender equity + mathematics, gender equity + STEM, education, mathematics + gender, mathematics + gender + America, mathematics + gender + education, mathematics + gender + United States military children, mathematics + military + children, mathematics + military + education, mathematics +*

military children + education, military + children, military + children + education, military children, military children + education, military children + schools, military +children + mathematics, military dependents + mathematics, military families, military families + education, military families + schools, military families + stressors, military personnel + families, military personnel + children, military personnel + youth, military youth, military youth + education, Nations' Report Card + mathematics + 2014+ Department of Defense, Sax + gender, Sax + gender + mathematics, STEM + gender equity, students + schools + armed forces, Terra Nova, Terra Nova + DoDEA, Terra Nova + military-connected students; and *high achieving military-connected students + STEM.* A particular challenge that I faced in conducting this search was locating current research relating to American high achieving military-connected students, mathematics achievement, and gender.

Theoretical Framework

In four separate publications, Sax (2005, 2009, 2010, 2011) conducted a meta-analysis of research on gender differences related to learning. From these meta-analyses, Sax presented key findings as well as recommendations about how to improve educational experiences for male and female students based on these findings. Sax's research in relation to these four publications formed the theoretical framework for this study.

Why Gender Matters

In an exploration of why gender matters and what parents and teachers need to know about the emerging science of sex differences, Sax (2005) noted that researchers have found evidence that preprogrammed differences exist in male and female brains (Achiron, Lipitz, & Achiron, 2001; Hanlon, Thatcher, & Cline, 1999). According to this research, male and female students use different portions of their brains for hearing, vision, language processing, numeracy, and problem solving. These differences influence their learning styles, which coupled with different maturation rates may contribute to disruptive classroom and learning environment behaviors for the respective genders. Not only do male and female students use different portions of their brains for hearing and vision, Achiron, Lipitz, & Achiron (2001) and Hanlon, Thatcher & Cline (1999) contended they do so in relation to learning styles as well. Based on this research, Sax argued that female students use both sides of the brain to process language, while male students use only the left side of the brain to process language. Based on this research, Sax also contended that male students activate a different portion of their brain than female students to solve arithmetic computations. Sax also noted that female and male children exhibit differences in the rods and cones of their eyes, which may account for why female children pay more attention to texture, color, and details, while male children are more attentive to actions.

Brain differences may also influence differences in behavior. Sax (2005) cited researchers who have found that male children play more aggressively than female children, even across primate species boundaries, generally preferring play involving actions like running and throwing, while female children are likely to engage in passive play such as pretend babysitting (Berenbaum & Snyder, 1995). Differences in learning styles surface early, Sax noted, often by school age. Young female students are more likely to attend to homework even when it is of little interest to them, if only to please their teachers. Male students, on the other hand, need to find some interest in the homework topic in order to actively engage with it.

Sax (2005, 2009, 2010, 2011) presented several key recommendations for improving instruction based on his findings that biological differences influence behavioral differences which in turn could influence educational success or failure. Sax's key recommendations included moving male students to the front of the classroom, differentiating instruction to meet these gender needs, and offering single gender education classes or schools. Because these behavioral and learning styles develop at staggered times in early childhood, Sax (2005, 2009) recommended that parents delay enrolling their children in school until they are 7 years old, especially boys. Sax (2009) recommended finding a more amenable balance between active experiential learning and scholarly or book learning for all children in their earliest school years. Sax also recommended delaying rigorous literacy activities such as reading and writing and numeracy curricula until first or second grade. Sax strongly advocated for single-sex education, not only within classrooms, but also at the school level. Because gender differences are significant in early childhood, Sax believed that single-sex education should be an option in elementary school classrooms.

Boys Adrift

In an investigation into a lack of academic motivation and achievement in young men, Sax (2009) found that five factors contribute to this problem. These factors include (a) instructional changes at school, (b) video games, (c) medications for attention deficit hyperactivity disorder (ADHD), (d) endocrine disrupters, and (e) lack of appropriate adult male role models. In relation to changes in instruction, Sax noted that one of the most striking changes in modern schooling is that rigorous literacy and numeracy curricula have now extended to the kindergarten level so that kindergarten is no longer a time for students to learn basic socialization skills. Instead, kindergarten students are required to sit still for long periods of time and complete curricula that in the past would have been delegated to older students. Sax maintained that this change is not developmentally appropriate, especially for young male students who have not mentally matured to benefit from complex curricula at such a young age. In contrast, Sax noted that the developmental trajectories of female kindergarten students are often more advanced than the development trajectories of kindergarten male students. Sax contended that the chasm between the brain development of male and female students seems to diminish around age 14. Therefore, Sax recommended that all children, and especially male children, should wait to start school until the age of 7.

Another factor that Sax (2009) contended impacts the academic motivation and achievement of male students is playing video games, which often contributes to a lack of time spent socializing or studying. Sax noted that young men spend twice as much time playing video games as young women. In addition, Sax noted that male students who are unmotivated in school are often highly motivated to play video games. Sax suggested that video games may affect the brain by throwing parts of the neural system that contributes to drive and motivation out of balance. This effect in the brain, Sax noted, is similar to the effect caused by psychiatric medications. Sax concluded that fascination with video games is often a contributing factor to derailing male students in achieving academic and social success.

An additional factor that Sax (2009) believed impacts the academic motivation and achievement of male students is endocrine disruptors. The endocrine system, Sax noted, consists of glands that produce hormones to regulate metabolism, growth, development, and other important body functions. Some human-designed chemicals imitate hormones and have adverse effects. These chemicals pervade the environment and can be found in seemingly benign objects such as baby bottles, plastic drink bottles, the linings of canned foods, and even in drinking water. The phthalates found in clear plastic drink bottles may negatively affect male reproductive systems. Sax noted that endocrine disruptors that accelerate puberty in young women often have the opposite effect on the sexual development of young men. Sax contended that increased rates of ADHD, increased rates of obesity, and the diminished drive to achieve in male students could be attributed to endocrine disruptors acting as environmental estrogens.

Another factor is the lack of role models for male students. According to Sax (2009), cultures are often defined by how they define manhood. Sax contended that being a man means “using your strength in the service of others” (p. 181) and learning this truth is accomplished by active experiential learning or knowledge gained through experience. Although definitions for manhood vary, Sax contended that young men in America are not likely to seek jobs they judge to be boring, uninteresting, or even beneath them. In addition, Sax believed that male students need more appropriate male role models to help them learn proper behavior, improve motivation, and develop skills in order to transition into adulthood. Sax defined adulthood as the ability to be independent of one's parents. Sax argued that young men often continue to live at home because of a lack of motivation to become economically or spiritually independent adults. Sax also noted that violent crimes committed by young men have increased in recent decades.

Sax (2009) attributed this lack of motivation and achievement to a collective failure of parents and society to assist young men in their transition to adulthood, particularly in providing effective adult male role models. According to Sax, no enduring cultures can be found where young men learn the rules of acceptable social behavior exclusively from women. To that end, Sax suggested that male children, including teenagers, should be enculturated by adults and not by their peers. Sax recommended that parents limit their male children's daily exposure to video games, restrict their male children from playing violent video games, and expose their male children to real-life experiences to counterbalance the simulated learning presented in video games and on television screens. Sax also recommended that parents and doctors carefully monitor medications because they might not be necessary or they might do more harm than benefit. In addition, Sax noted that ADHD is not a new disease. By 2009, Sax noted, children in the United States received psychiatric medications three times more frequently than European children. Sax contended that these medications might damage the developing brains of male children, which could affect their motivation to learn and perhaps lead to feelings of apathy and disengagement. Sax also recommended eliminating all other possible causes for male children's disengagement from learning before diagnosing the disorder as ADHD and prescribing potentially harmful medications. Sax believed that enrolling male students in single-sex schools might produce the same ameliorating effects as medications. As Sax (2009) suggested, “We should not medicate boys so they fit the school; we should change the school to fit the boy” (p. 96). Sax also suggested that adults should guide young men into gender-separate communities that reinforce social skills necessary for them to become self-actuated adults, such as the Boy Scouts, competitive sports, the Isaac Walton League, or other outdoor-oriented organizations with adult supervision. In addition, Sax suggested that single-sex schooling may be a better format than coed schooling to prepare young men for interacting in the real world.

Girls on the Edge

In an investigation of the problems that young American females face in modern society, Sax (2011) noted that research indicates the following four factors contribute to the crises that young females in American society face: (a) sexual identity, (b) obsessions with online activities and personas, (c) other obsessions such as obsessions with body image that could lead to eating disorders including anorexia or bulimia, and (d) environmental toxins (Force, 2007; Martin, 2007; Sathyanarayana et al., 2008; Silverthorne, 2009;Steingraber, 2007; Wichstrom, 1999). In the area of sexual identity, Sax discussed the sexualization of young American girls through media and modern culture. Sax noted that girls in the last few decades have been encouraged to wear sexualized clothing before puberty. Sax contended that placing prepubescent girls in sexualized clothing is unhealthy because they become objects on display for others, and this type of dress sets them up for depression, anxiety, and potentially unsatisfying sexual relationships later in life. Sax believed that modern American society pushes young girls to adopt sexual identities and to become sexual agents and sexual objects much too soon. Sax argued that young girls should be treated as young girls for as long as possible before graduating into womanhood.

Concerning online activities, Sax (2011) addressed the *cyberbubble*, which he defined as young people obsessed with online activities and online personas. The almost universal availability of social interactions via the internet has created literal networks of social interactions, which means that girls will often create social personas that may or may not be genuine. Sax referenced research that Bauerlein (2008) conducted about social networking by teenagers, including creating and maintaining online personas and research on the intellectual life of young adults and the possible negative impact of on line relationships to American culture. Sax noted that Bauerlein's research suggested excessive online hours may have a statistically negative relationship with mathematics and reading skills and that skills acquired and honed online might actually impair a person's later performance in adult workplaces. Bauerlein concluded teenage use of simplistic syntax, phonetic spelling, and poor diction could create poor communication habits. Sax was concerned that the cyberbubble could also become a trap where pre-teens and teenagers pay so much attention to their cyberpersonas that they might fail to find a sense of place or a sense of who they really are.

In relation to other obsessions, Sax (2011) maintained that young females who lack a healthy sense of self could become vulnerable to unhealthy obsessions, such as anorexia, excessive athleticism, overuse of prescription drugs, use of illegal drugs, alcohol abuse, and emphasis on academic achievement to the exclusion of other activities. Sax also noted that addressing these obsessions might be the crucial step in preventing girls from hurting themselves through negative behaviors and promoting the development of a healthy sense of self. According to Sax, adults need to create environments where young girls participate in age appropriate behaviors. Sax believed that adults should not encourage, or even tolerate, children participating in activities that are typically reserved for adults.

Concerning environmental toxins, Sax (2011) expressed concern that many American female students reach puberty at younger ages than in previous decades, which may be due to over exposure to environmental toxins. Sax noted that humans have the longest childhood of any mammals and their ability to learn completely new skills (e.g. language and walking) before puberty diminishes profoundly after puberty. Sax referenced research that Steingraber (2007) conducted about the decreasing age of puberty. Steingraber noted that many American girls reach puberty when they are 10 years old. Sax identified several environmental toxins that may influence the early onset of puberty and emphasized how reducing the number of years a female spends in childhood could be detrimental to her developing physical and emotional health. Sax also noted American females have been exposed to other environmental toxins such as bisphenol A that can be found in hard plastic food containers, phthalates that are found in lotions and skin creams, and polyethylene terephthalate ethylene that is also a main ingredient in clear soft plastic bottles.

To address these factors, Sax (2011) recommended that parents should limit their daughters' exposure to harmful environmental toxins such as chemicals in some plastics that might cause early-onset puberty. Sax also recommended that parents should become aware of their daughters' exposure to cultural and societal pressures such as early sexualization through commercialization as well as social interactions, especially via the internet. In addition, Sax suggested that parents should understand how their daughters interact in online environments in order to ensure their daughters do not participate in negative online relationships that could hurt their self-esteem and that parents should also reduce online activities that could impair their daughters' performance in school or in future work environments. Sax recommended that parents should cultivate their daughter's self-esteem and spiritual beliefs to promote healthy behaviors through adolescence into adulthood and to prevent unhealthy obsessions or depression. Sax concluded that the best practices for teaching young girls may be fundamentally different from best practices for teaching young boys. Therefore, Sax recommended gender-specific instruction as well as single-gender classrooms or schools so that education could be directed towards the learning needs of female students.

Gender Differences in Hearing

Sax (2010, 2016) also discussed educational implications related to gender hearing differences. In an earlier publication, Sax (2010) stated that male children have a higher tolerance for noise than female children. In a more recent online update, Sax (2016) reinforced his earlier contention that young male children tolerate loudness in the classroom better than young female children. However, Sax (2016) noted that newer research that Sagi, D'Alessandro, and Norwich (2007) conducted indicates that this difference is probably only three decibels, which might not be relevant in a normal classroom. However, Sax contended that the notion that male and female students may have different comfort levels related to noise levels in the classroom could still support the need for single gender classrooms or the use of selective amplification devices. Sax (2010) recommended that teachers place male students at the front of the classroom so that they can hear the teacher, especially if the teacher is soft-spoken. Sax also contended that teachers need to be attuned to noise levels in the classroom because female students may have a lower tolerance to noise and may have difficulty concentrating in a noisy classroom or may feel that a male teacher is shouting at them when the teacher believes he is speaking in a normal tone.

In summary, scientifically verified physical and emotional differences between the genders, especially through the age of puberty, are the bedrock of Sax's (2005, 2009, 2010, 2011, 2016) arguments for gender-based educational options. Sax (2009) recommended differentiating instruction to meet gender differences among students. Sax suggested that teachers include activity in their instruction because male students may be more attentive to movement due to differences in specialized photoreceptor rod cells in their eyes that are sensitive to light and dark changes as well as shape and movement and cone cells that are sensitive to color. In a similar context, Sax recommended increasing hands-on learning activities to help male students become more actively involved in learning. Sax (2010) also recommended increasing opportunities for cooperative learning strategies for female students and competitive learning strategies for male students and that these strategies be tailored to benefit the majority of students in each gender. Sax believed that teachers should not submit young females to learning experiences that are best tailored to the strengths of young male students and vice versa. Sax (2005) advocated for single sex education because he believed that gender neutral education may favor one gender over the other, leading to the selection of stereotypical careers by male and female students. Another key recommendation from Sax (2005, 2009, 2011) was that educators should teach to different learning styles in order to help all students become confident enough about their own abilities to pursue the education and careers they desire. Sax (2005) also contended that male students often play more aggressively than female students, and teachers should understand and recognize those differences, perhaps even incorporating them into learning activities, instead of disciplining or medicating male students for typical male behavior.

Articulation in Current Research

Other researchers have also articulated Sax's research on gender-based instructional practices. Tichenor, Welsh, Corcoran, Piechura, and Heins (2016) surveyed 168 females in both mixed-gender and single-gender classrooms in Grades 1 through 5 to discover student attitudes towards mathematics. Tichenor et al. found that female students in single-gender classrooms applied mathematics to real-life situations more frequently than students in mixed-gender classrooms. Tichenor et al. concluded that female students in single-gender classroom settings are more likely to understand the usefulness of mathematics and how mathematics can be applied to real-life situations. In addition, Tichenor et al. concluded that female students in single-gender classrooms were approximately half as likely as female students in mixed-gender classrooms to perceive mathematics as a boring subject.

In related research about gender differences in learning, Arroyo, Burleson, Tai, Muldner, and Woolf (2013) compared the responses of 544 high school students in Grades 9 and 10 to advanced adaptive technology in a computerized mathematics program that presented students with gendered learning companions (e.g. male or female animated tutors). Arroyo et al. found that female students benefited more from animated learning companions, especially female companions, than male students. Arroyo et al. suggested that the gender of animated learning companions should be tailored to the gender of the students for best results.

In a multiple case study, Kao (2015) examined nine mathematically gifted female Taiwanese students' perceptions of gender stereotypes. Kao found that these middle school students disliked gender inequity, noting that they "agreed with the stereotypes of negative female dispositions, identified with masculine traits, and disliked all-girl environments" (p. 25). Kao concluded that because these students belonged to multiple groups and subgroups, this factor often led to complex belief systems and perspectives. Kao suggested that in order to maintain gender equity in society, mathematically gifted females should be encouraged to value and identify with positive feminine characteristics.

In a study of stereotype threat, Mangels, Good, Whiteman, Maniscalco, and Dweck (2012) used electroencephalogram (EEG) monitoring to focus on the effects of emotion during mathematics testing, comparing those students who rebounded from mathematics errors to those students who did not. Seventy-one female undergraduate students were divided into two groups exposed to either stereotype threat or non-stereotype threat in relation to testing conditions. Results showed that stereotype threat had a negative impact on testing results for female students. Mangels et al. called for more research to understand how to provide more effective learning for female students in mathematics classes.

In a different study of stereotype threat in mathematics, Casad, Hale, & Wachs. (2017) compared female students in regular and honors classes to determine if gender identity was a risk factor. Casad et al. examined attitudes towards mathematics, disengagement, and math performance. Casad et al. defined disengagement as having no connection between success or failures in mathematics and feelings of self-worth. A female who disengages would not have lower self-esteem if she did poorly on a mathematics assessment or in a mathematics class. Casad et al. noted that chronic disengagement could cause females to leave STEM fields. Casad et al. defined gender identity as the importance of gender to feelings of self-worth. Casad et al. stated females with strong gender identity are more likely to be influenced by the negative stereotype threat that males are inherently better at mathematics than females. Females might be overly concerned that potential poor performance on their part would reflect on the female population as a whole. Casad et al. found differences between the female students in regular classes and the high achieving students who chose to participate in honors classes. factor for females in the regular classes, but protective for honors students. However, gender identity in regards to female middle school honors students was a risk factor for actual mathematics performance but a buffer for the students in the regular classes. Casad et al. suggested more research should be conducted in the areas of stereotype threat in regards to specific educational contexts.

Similarly, Tine and Gotlieb (2013) examined the effect of stereotype threat on mathematics performance, going further to categorize the threat as gender-based, race-based, income-based, or multiple-minority based. Tine and Gotlieb administered mathematics tests to 71 undergraduate students between the ages of 18-26. Tine and Gotlieb found stereotype threat effects were not present for the variable of gender alone but were present for the variables of gender and working memory together. In addition, Tine and Gotlieb found the effects for income-based stereotype threat were the strongest for mathematic performance when examining single aspects of identity. However, Tine and Gotlieb concluded that both mathematics performance and working memory are more likely to be negatively affected by the presence of stereotype threat in persons with three stigmatized aspects of identity than in persons with less than three stigmatized aspects of identity. Tine and Gotlieb recommended that educators be mindful of the confounding effects of students' total identity compositions.

Watt et al. (2012) studied gender differences in mathematics-related career aspirations. A total of 1,247 students from Australia, Canada, and the United States were involved in their study. Although mathematics courses were compulsory in all three countries for students in the early secondary grades, students in the upper grades were generally allowed to choose whether or not to continue with mathematics coursework. Watt et al. found that student success in mathematics courses relative to career aspirations was important to Australian and Canadian students, but not to United States students. Watt et al. concluded that requiring students to take higher level mathematics courses does not translate into higher educational or occupational aspirations, and in this regard, Watt et al. found no significant differences in the genders.

This study benefits from Sax's (2005, 2009, 2010, 2011) findings on gender differences related to learning because these findings confirm that gender differences in academic environments are real and identifiable. In addition, this study benefits from Sax's previous research findings that gender differences may contribute to observable differences in academic achievement. Finally, this study benefits from Sax's findings that gender differences can be addressed in positive ways to reduce or eliminate gender bias. Concerning the potential findings for this study, it is possible that mathematics scores on the TNTE might show disparate success for male and female students. That, in itself, is unlikely to explain why gender differences do or do not exist. However, because male and female students may learn differently, if differences in mathematics scores do exist, those differences may be a stimulus to discover the causes.

Literature Review

The relationship between gender and academic success in STEM fields has been widely researched. However, the relationship between gender and success in mathematics, specifically for military-connected children, is under-reported in current research. Gender has also not been well established as a reliable predictor of success or failure in mathematics. Grade attainment and test scores are variables that have been considered relative to mathematical achievement. To this end, variables such as instruction, assessment, and gender in relation to learning mathematics are viable avenues to investigate. Therefore, this literature review is organized into three major sections. The first section includes an analysis of current research about gender differences in mathematical learning. The second section includes an analysis of current research about gender differences in mathematics achievement as measured by large scale international and national assessments for K-12 students in both schools in the United States and other countries. The third section includes an analysis of current research about mathematics achievement for military-affiliated children.

Gender Differences in Mathematical Learning

A gender gap in mathematics has been documented for decades (Hyde, Fennema, & Lamon, 1990; Tyre, 2008; Watt et al., 2012). In earlier research, Hyde, Fennema, and Lamon (1990) conducted a meta-analysis of 100 studies relating to gender differences in mathematics performance over a 15-year time period, which "represented the testing of 3,175,188" subjects (p. 139). Hyde et al. found that female students scored slightly better than male students when they averaged all of the effect sizes over the general population. An analysis of age trends found no gender differences in the area of problem solving but female students outperformed male students in computation at both the elementary school and the middle school levels. However, at the high school level and college levels, the trend reversed and male students outperformed female students. In an analysis of the highest achieving students, Hyde et al. found that the gender gap was the largest. They also found that the gender gap was larger in studies "published in 1973 or earlier" (p. 139). Hyde et al. concluded that although the gender gap had shrunk since 1973, and although the gap in mathematical performance was small for most groups, the fact that female students performed lower than male students in problem solving should be of concern to future researchers. This study is important because although more female students graduate from college than male students, fewer female students than male students are employed in STEM careers, some of which are dependent on mathematics abilities or skills.

In earlier research, Tyre (2008) examined mathematics scores on the NAEP for 17-year-old students in both public and nonpublic United States schools from 1973 to 2004. Tyre found that from 1973 through the 2004 NAEP testing cycles, 17-year-old female students scored lower than 17-year-old male students. Tyre reported that the gender gap favoring male students decreased from an 8-point difference in 1973 to a 3-point difference in 2004. Tyre concluded that male and female students may perceive their mathematical abilities differently, which might result in different attitudes about mathematics and different approaches to teaching mathematics. Tyre also concluded that male and female students may be subject to gender stereotyping, albeit to different degrees, and subject to social and psychological stresses that could lead to anxieties affecting their performance and mathematical successes. Therefore, the following subsections include an analysis of current research related to student perspectives and attitudes about mathematics and mathematical abilities, gender stereotyping, and psychological stressors.

Student Perspectives and Attitudes About Mathematics and Mathematical Abilities

Ganley and Lubienski (2016) examined gender-related patterns of achievement for 7,040 students in Grades 3 through 8 over a 5-year period. Ganley and Lubienski focused on differences in mathematics confidence, interest in mathematics, and mathematics performance. Ganley and Lubienski found that male and female students became less confident and less interested in mathematics while progressing from elementary to middle school. They also found that gender differences in mathematics confidence and performance were smaller in the eighth grade than in the earlier grades. Ganley and Lubienski found female students to be less confident and less interested in mathematics than male students. Because the gender gap in performance was less than the confidence and interest gaps, Ganley and Lubienski concluded that the lack of confidence by female students was not warranted. Ganley and Lubienski also found that the best predictors of confidence and interest in mathematics were the earlier degrees of confidence and interest in mathematics that these students demonstrated. Ganley and Lubienski concluded that student interest and confidence in mathematics are malleable and suggested that the small gender differences in the earlier grades could be exacerbated later. Consequently, Ganley and Lubienski suggested that early interventions, especially for female students, should begin in the elementary grades. This study is important because results demonstrated a difference in mathematical confidence and interest as students grow older. Ganley and Lubienski also referenced similar studies in other countries and addressed the relationship between reciprocal variables impacting student success in mathematics, which means that inferences can be drawn relative to potential interventions to promote gender equity in mathematics.

In other related research, Grunspan et al. (2016) compared perspectives of male and female students about their perceived knowledge of course content in science. The sample included 1,715 students enrolled in three sections of a college biology course in a major university in the United States. Classes A and B employed male instructors, and Class C employed two male instructors and one female instructor. In all three classes, female students outnumbered male students, but only slightly. In all three classes, male students earned, on average, higher grades. Students were asked to anonymously predict peers who would do well in the course. Students were also polled again following three mid-course tests, accounting for 11 peer perception surveys. Students receiving the most nominations were labeled as celebrity students. Course instructors were asked to rate students according to how outspoken students were, although potential instructor bias was not addressed. Male students consistently received more nominations than female students, and this number increased as the course progressed. Female students in Class C received more nominations than male students. In all 11 surveys, male students exhibited a significant bias towards nominating male celebrities. Overall, female students showed no bias in nominating male or female celebrities, except that female students in Class C nominated more male celebrities on the final survey. Performance and outspokenness were correlated with nominations, and Grunspan et al. hypothesized that a positive correlation might account for nomination status. Even though male students nominated more male students over the span of the course, female students showed no such bias.

Grunspan et al. concluded, therefore, that male students are biased against their female peers. Grunspan et al. also found that female students in every class who performed successfully and were outspoken were not nominated as frequently as male students. Course instructors overwhelmingly rated male students as more outspoken than female students, although outspoken female students with grades equivalent to their male counterparts were found in every class. Grunspan et al. also concluded that maleness was a prerequisite for nomination. If male peers and male instructors continue to exhibit gender bias in STEM courses, Grunspan contended, it follows that gender disparities may persist in retaining female students in STEM courses and encouraging female students to enter STEM careers. This study is significant because it highlights gender bias as an underlying and persistent cause of gender imbalance in STEM career fields.

In a study about gender differences in value beliefs about mathematics, Gaspard et al. (2015) begin with the premise that male students are overrepresented in STEM fields related to mathematics. Gaspard et al. collected data from 1,867 German students approximately 14 years old by using questionnaires that trained research assistants administered. Slightly more females (53%) than males were surveyed. One goal of this study was to investigate gender differences in value beliefs relative to mathematics. Gaspard et al. found the structure of value beliefs to be similar for male and female students, albeit male students showed considerably higher positive values (i.e. personal importance) for mathematics, while females perceived higher costs (i.e. more effort required, negative emotions, and the opportunity cost of forfeiting time from other activities). Therefore, Gaspard et al. concluded that female students believed mathematics to be less important and less personally useful for future personal goals than male students. An exception was that female students believed mathematics to be more useful for school (i.e. grades) than male students. Gaspard et al. also found that gender differences related to value beliefs continued to be larger than gender differences in mathematics achievement. Gaspard et al. suggested that teachers provide interventions designed to increase positive attitudes towards mathematics for female students in order to shrink the gender gap in mathematics performance. This research is important because the results support the consideration of value beliefs when investigating gender differences in mathematics learning. If interventions are needed to reduce gender inequities in the perceived value of mathematics, future researchers might consider Gaspard et al.'s study to create such interventions.

In a related examination of gender and socioeconomic differences, Guo, Marsh, Parker, Morin, and Yeung (2015) conducted a multi-cohort study on predictors of mathematics achievement and aspirations. The change in Hong Kong's sovereignty from the United Kingdom to China in 1997 provided motivation for Guo et al.'s study. The target population was grade eight students in Hong Kong who had participated in the Trends in International Mathematics and Science Study (TIMSS) tests for 1999, 2003, and 2007. The multiplicative relations of expectancy and value on outcome variables comprised the main target of their investigation. However, Guo et al. also examined the background variables of gender and social economic status (SES) and their relationship to predicting mathematics achievement and mathematics-related aspirations. Guo et al. predicted that mathematics self-concept would be a strong predictor of mathematics achievement and utility value would be a strong predictor of aspiration. Guo et al. also conducted a gender analysis. The sample included 13, 621 students, evenly split between genders, with an average age of 14.4 years. As Guo et al. predicted, male students exhibited higher self-concept, which may indirectly lead to high mathematics achievement. However, when Guo et al. took into consideration self-concept and the intrinsic value of mathematics, female students demonstrated higher mathematics achievement than male students. Consequently, Guo et al. found no overall gender differences in mathematics achievement. As Guo et al. expected, educational aspirations favored female students to a small extent in their study. However, socioeconomic status exerted a bigger influence on the aspirations of male students than female students. Yet, despite only negligible gender differences in mathematics achievements, Guo et al. found a pattern of gender stereotyping favoring male students. Guo et al. concluded that the positive interactive effects between expectancy and values and the negligible gender neutralizing effects they produced may not generalize to Western cultures, or even other Asian countries. This study is important because improving student motivation may promote improved student achievement, regardless of socioeconomic status. If a comparison can be made between Asian and Western educational cultures relative to the TIMSS results, Guo et al.'s findings might imply that differences between male and female student achievement in mathematics is negligible.

In another significant gender study, Pahlke, Hyde, and Allison (2014) conducted a meta-analysis of the effects of single-sex schooling compared to coeducational schooling on students' performance and attitudes. Their study encompassed 184 studies, involving

1.6 million students from the United States and 20 other nations. Because single-sex schooling has been commonplace in other parts of the world for a long time and is only a relatively recent phenomenon in the United States, Pahlke et al. reported statistical data from United States schools separately with the intent of providing a clear interpretation of this data. Even though uncontrolled studies that Pahlke et al. reviewed revealed modest advantages for single-sex schooling in mathematics performance for both male and female students, only insignificant differences were found in the controlled studies. In some cases, small differences between single-sex and coeducational schooling actually favored students in coeducational schools. After analyzing the data from 21 countries, Pahlke et al. found that single-sex schooling did not provide an advantage to students according to dosage (i.e., single-class instruction versus whole-school instruction). When single sex schooling did provide a small advantage, it was in smaller dosages (i.e., single class instruction). Pahlke et al. also found that single-sex schooling did not provide an advantage for ethnic minorities, but they noted this finding may need further exploration. Unable to ascertain whether or not single-sex schooling reinforces gender stereotyping, Pahlke et al. concluded that single-sex schooling did not provide an advantage for ethnic minorities or any substantial advantage for students of either gender. Pahlke et al. was also unable to ascertain whether or not single-sex schooling reinforces gender stereotyping. Pahlke et al. had hypothesized that gender stereotyping would be particularly high in single-sex schools. However, Pahlke et al. concluded that not enough data were available to determine if gender stereotyping was impacted by a school's stakeholders' underlying assumption that essential (i.e. biological) differences affecting learning exist between genders. Such a guiding assumption could be transmitted to teachers, parents, and students in a single-sex school system. Pahlke et al. coded for such assumptions in only 8% of the schools they studied. In order to draw meaningful connections between gender stereotyping and gender-related messages exchanged between citizens of a single-sex school system, Pahlke et al. believed that future researchers would have to first determine a school's basic and underlying assumptions relative to biological differences between stakeholders. Pahlke et al. also recommended the need for further research about single-sex schooling for ethnic minorities, especially for male African American students because few studies exist that address the efficacy of single-sex schooling for these groups. This research is important because gender stereotyping can play a significant role in mathematics education for female students.

In an earlier study about gender and mathematics, Mendick (2005) claimed that male students approach mathematics differently than female students. Mendick asserted that male students relate to mathematics through an approach of separateness. In contrast, Mendick claimed female students relate to mathematics through an approach of connectedness. Mendick hypothesized that students' understanding of mathematics would improve if teachers moved away from individual, abstract, rational, and objective (masculine) ways of teaching mathematics and towards relational, grounded, emotional, and subjective (feminine) ways of teaching mathematics. To support that hypothesis, Mendick interviewed 43 students in London who elected to continue with mathematics coursework beyond compulsory schooling that ended at age 16. From interviews and observations, Mendick constructed two psychoanalytic stories (one about a male student and one about a female student) to support the finding that teachers need to consider how educational systems reinforce gender biases. This study is important because learning mathematics from objective (masculine) and subjective (feminine) viewpoints supports the notion that male and female students may learn mathematics differently.

Gender Stereotyping

In a significant study about gender stereotyping, Casad, Hale, and Wachs (2015) studied social determinants for mathematics anxiety in adolescents. Their two-part study looked at the role of parents' mathematics anxieties relative to their children's mathematics anxieties and the role it plays in mathematics performance. Casad et al. examined the ways in which mathematics-gender stereotypes could predict mathematics anxieties and outcomes. In the first part of this study, 683 parents participated with their children (55% female, 45% male) aged 11 to 14 who were enrolled in Grades 6, 7, and 8. For the second part of the study, researchers administered questionnaires to 1,342 students (53% female, 47% male) addressing their beliefs about gender stereotypes related to mathematics. Casad et al. found that gender stereotypes have negative effects on both female and male students. One unexpected finding was that when mathematics anxiety for fathers was high but sons' mathematics anxiety was low, their sons' grade-point averages were high. Casad et al. also found that female students were more negatively impacted by cultural bias than male students. This study is important because it supports the contention that socialization of male and female students is impacted by the beliefs, behaviors, and expectations of parents and teachers. This research also supports the importance of parent involvement in their children's education.

In an examination of factors that might predict student achievement and behavior in undergraduate mathematics coursework, Alcock, Attridge, Kenny, and Inglis (2014) equated personality with five major dimensions, including conscientiousness, extraversion, agreeableness, neuroticism, and openness to experience. Alcock et al. maintained that gender is a social construct, which is different from biological sex inheritance, and that gender roles are socially constructed. The gender of participants was deduced from the self-selection of preferred titles (i.e., Mr./Mrs./Miss/Ms.), and it was assumed that gender would coincide with the biological sex of participants. Alcock et al. maintained that because men dominate enrollment in STEM courses and in some STEM career fields, gender imbalances should be investigated from a social perspective. Alcock et al. targeted 89 undergraduate students in a United Kingdom (UK) school of mathematics. Students completed a self-report at the beginning of the study, which was intended to reveal aspects of their personality and their approaches to learning. Alcock et al. analyzed course data with respect to gender first. Then they controlled for personality and compared the predictive amount of academic success variance by gender and by personality. Alcock et al. found that personality has a greater predictive power than gender in the case of mathematical achievement. This research has important social implications because the results from this study, like the results of the Grunspan et al. (2016) study, confirm that female students in some STEM courses are part of a disadvantaged group due to widespread stereotyping of mathematics as a male-dominated endeavor. The important conclusion to draw from Alcock et al.'s (2014) research is that factors other than gender should also be considered as more accurately predictive of academic success in STEM coursework. Specifically, Alcock et al. concluded that an individual's personality, rather than gender, might more accurately predict success in STEM courses and ultimately STEM careers. From this perspective, Alcock et al. believed that it might be more productive to increase female inclusion in STEM education and careers by de-emphasizing female students as part of a disadvantaged homogenous group and, instead, publicizing data showing that female students achieve academic success equal to, if not superior to, male students in STEM coursework.

In related research, Picho, Rodriguez, and Finnie (2013) conducted a meta-analysis to investigate the role of context in relation to mathematics performance of female students under stereotype threat. They distilled their original search for data to 103 studies found in 44 articles and dissertations for a 17-year period from 1995 to 2011. Their study addressed the following three hypotheses relative to female students in testing environments: (a) implicit priming of subjects would produce larger negative effects than would explicit priming, (b) more negative effects would be produced where female students were in the minority and smaller stereotype threat effects would be produced where female students formed the majority, and (c) smaller stereotype threat effects would be produced in the northeastern region than in the southern region of the United States. Participants included a total of 5,588 female students with 2,820 of these students assigned to control groups. Picho et al. discovered explicit (overt) priming produced more variance in test results than did implicit (covert) priming, but the difference in effect size was not statistically significant. When they compared mixed gender to single gender samples, Picho et al. discovered that stereotype threat effects produced larger effect size variance in the mixed gender samples. However, the differences were not large enough to be statistically significant. Picho et al. also found no statistically significant differences in effect size due to the regional locations of subjects, although they did find more variability in the northeastern and western regions of the United States than in other regions. Picho et al. also found that placing female students in single-gender or female-majority testing environments did not improve test scores. Picho et al. concluded that stereotype threat exerts a negative effect on the testing performance of female students, but few replication studies are available that investigate the multitude of moderating factors that may contribute to driving the negative effects of stereotype stress on females, especially in STEM fields. Picho et al. also concluded that priming, the sex composition of classes (mixed-gender versus single-gender), and the region of the United States do not significantly moderate females' mathematics performance. In addition, Picho et al. concluded that a gap exists in current research about stereotype threat and female mathematics performance. They claimed that too little data are available to advocate the effects of stereotype threat on female students in mathematical testing environments. However, Picho et al. also found that female students who completed mathematics tests under stereotype threat conditions scored lower than female students in control groups. Picho et al. also found that level of education and geographic location impacted student performance in mathematics. In addition, they found that stereotype threat negatively affected mathematics performance for female students in middle school and in high school more significantly than for female students in college. They also found that females who completed tests under stereotype threat conditions and who lived in countries with less notable gender bias and small gender gaps (e.g. Netherlands and Sweden) produced test results with little variances in effect size while female students who completed tests under stereotype threat conditions in countries with more notable gender biases (e.g. African countries) had larger effect sizes. This study is relevant because it confirms the notion that variances exist between male and female students in relation to mathematics test results and affirms contributing or moderating factors that are not fully understood. Picho et al. also noted a gap in the literature relating to female students and mathematics testing and replicable studies on stereotype threat and mathematics. Their study also showed a difference in how stereotype threat affects students from different educational levels. Because military-connected students are part of an identifiable military subculture, components of Picho et al.'s (2013) study might suggest areas of further research.

In another study about gender stereotyping, Johnson et al. (2012) examined the effects of stereotype threat (ST) and stereotype lift on males' and females' mathematics test performance. Johnson et al. asked three questions relative to mathematics test performance, all controlled for past stereotype threat vulnerability: (a) What is the effect of gender? (b) What is the effect of stereotyping condition of threat, lift, or neither? and (c) What is the effect of the interaction of gender and stereotyping condition? The sample consisted of 458 volunteers (178 males, 280 females) aged 18 to 55. All participants were college students representing 24 different college majors and five ethnic groups. Prior to experimental testing, participants completed a *ST Vulnerability Scale.* Participants in the experimental group received ST ability-based prompts (e.g. men are expected to do better on this test or females are expected to do worse on this test) before testing. Johnson et al. found that male and female students react differently to stereotype threat and stereotype lift. Johnson et al. also found that male students performed better when under stereotype threat than under no threat or under stereotype life conditions while females performed better under no threat or under stereotype lift than under stereotype threat. That is, male students performed better under stereotype threat and female students performed better under stereotype life. From a practitioner's point of view, Johnson et al.'s results show how teachers might improve the mathematics test performance of female students by using stereotype life strategies such as reminding females that they belong to social groups other than gender groups (e.g. a group of intelligent students). Johnson et al. concluded that there is great benefit in building student confidence in mathematics early in their education in order to improve the participation of female students in STEM education and careers. This study is relevant because it affirms the notion that female students and male students respond differently to mathematics testing environments.

Early references to stereotype threat were directed at differences between members of different races and their performance on tests (Steele & Aronson, 1995). Soon after, the concept of stereotype threat was applied to the gap between male and female mathematics performance (Spencer, Steele, & Quinn, 1999). Stoet and Geary (2012) considered the Spencer et al. (1999) study as the original report dealing with stereotype threat and the gender gap in mathematics. Stoet and Geary noted that since this report, the hypothesis that members of any socially identifiable group might be subject to a self-fulfilling performance on mathematics tests based on perceived group characteristics has been widely accepted in both the popular media and academia. Stoet and Geary questioned whether or not published research provided enough evidence to support stereotype threat as the primary cause of differences between male and female performance on difficult mathematics tests. Stoet and Geary first identified 141 published articles relative to stereotype threat and gender gap in mathematics. Of those 141 articles, 23 articles met criteria establishing them as valid repetitions of the original research that Spencer et al. conducted in 1999. Stoet and Geary then conducted a meta-analysis of these 23 valid studies and found that the gender gap only surfaced in those studies that applied a moderator (a third variable that affects the correlation of two other variables) of previous mathematics performance to adjust scores. For studies that were unadjusted by such a moderator, the gender gap did not surface. Thus, Stoet and Geary concluded that not enough supporting data exists to affirm stereotype threat as the primary cause of gender differences on difficult mathematics tests. Other factors may come into play when attempting to ascertain the causal reasons for a mathematics gender gap. Stoet and Geary predicted that continued adherence to the stereotype threat hypothesis might hinder research towards other possible contributing factors for the gender gap in mathematics and ultimately in STEM careers. Because mathematics is a major component of STEM, it acts as a gateway into STEM careers, which prompts a concern about alleviating the gender gap in mathematics. According to Stoet and Geary, lack of research regarding alternative causal factors may negatively impact the development and implementation of interventions to reduce the gender gap in STEM.

This study is important because the results showed stereotype threat may not be a major cause of the gender gap in STEM careers.

In related research, Plant, Theoret, and Favreau (2009) challenged the trend in gender stereotyping that aligns maleness with mathematical prowess and femaleness with language superiority. Plante et al. conducted a study on gender stereotypes with 984 French-speaking Canadian students in Grades 6, 8, and 10. Plante et al. administered a modification of the *Mathematics as a Gendered Domain* that Leder and Forgasz had developed in 2002. Plante et al. hypothesized the following: (a) that students would view mathematics as equally suited to male and female students, (b) that students would view language as a more feminine domain, (c) that male and female students would both favor their own gender when expressing stereotype preferences, and (d) that student adherence to expressed stereotypes would weaken as they became older. Participating students were asked to rate questionnaire items using a 7-point Likert scale. Testing was divided into two 20 to 30-minute sessions. Half of the group began with the mathematics questionnaire. The other half began with the language questionnaire. *T*-test results showed that grade six male students perceived mathematics to be more in the male than the female domain, but by Grades 8 and 10, male students appeared to be neutral in that regard. Female students of all ages reported mathematics to be more in the female domain. No clear difference emerged regarding the effect of school level (i.e., grade level) on male or female stereotypes. However, a general tendency was found for both male and female students to favor female students in both the mathematics and language domains. This overall tendency is a reversal of traditionally perceived gender stereotypes. Addressing whether or not traditional stereotyping favors male students in mathematics, Plante et al. called for more research to discover if there is support for the contention that teenage students may now believe mathematics is within the female domain or is neutral relative to gender. This research is particularly relevant because peer stereotyping is an important variable to consider in gender stereotyping.

Psychological Stressors

Schommer-Aikens, Unruh, and Morphew (2015) examined mathematical anxiety and subsequent performance in vocational technology students. Schommer-Aikens et al. found that students come to vocational technology school with predetermined beliefs about their own mathematical skills, the difficulty of mathematics, the usefulness of mathematics, and how to solve difficult mathematical problems. Students who have different epistemological beliefs than their instructors are more likely to score poorly on mathematical assessments than students who have similar beliefs as their instructors. Similarly, students who demonstrate mathematical anxiety are more likely to score poorly on mathematical assessments than students who do not demonstrate mathematical anxiety. Schommer-Aikens et al. concluded that student beliefs are adequate predictors of both mathematics anxiety and mathematical performance. This study is important to gender-based research about mathematics because the correlation between mathematical anxiety and mathematical performance needs to be considered. If female students experience more mathematics anxiety than male students, anxiety could be a reason for a gender gap in mathematics.

Similarly, Spitzer and Aronson (2015) conducted an extensive literature review of several studies. Spitzer and Aronson (2015) summarized the results of studies on social psychological interventions such as mediation, role model exposure, and growth mindset to reduce educational disparities of adolescents. They found that female students lose confidence in STEM subjects as they get older and subsequently lose interest in taking advanced STEM courses. By reviewing multiple reports, Spitzer and Aronson addressed social psychological interventions used in the United States and attempted to assess their efficacy in reducing disparities in educational achievement gaps. They focused on social psychological interventions intended for use in altering the perceptions of stereotype threat and how students respond to it. Spitzer and Aronson contended that serious consequences result in the ways educators attend to educational gaps. The interventions Spitzer and Aronson reviewed included mediation, role model exposure, reappraisal, growth mindset, possible selves, values affirmations, belonging interventions, and cooperative learning. According to Spitzer and Aronson, achievement gaps are only partly explained by objective structural barriers (e.g. genetics, teachers, wide sweeping policies) and are impacted more by the subjective experiences of students. Although acknowledging that more research is needed about interventions, Spitzer and Aronson concluded that there is much to be optimistic about in empowering students with ways to address their psychological experiences with perceived inequalities (i.e. stereotype threats) rather than implementing expensive, nationwide programs of rewards and punishments. Providing students and teachers with techniques to allow them to make the best of their unequal opportunities may hold more promise in narrowing the academic gap than past and present policies. This research is important because the interventions Spitzer and Aronson discussed might help students cope with psychological stressors such as stereotype threat that may contribute to achievement gaps in mathematics.

Prompted by the poor performance of students in the United States on the PISA and by the poor performance of grade eight students on the NAEP, Chu, Guo, and Leighton (2014) sought to expand on a theoretical understanding of student performance on these standardized tests by exploring the affective variables of interpersonal trust and attitudes in relation to participating in standardized tests. Chu et al. defined interpersonal trust as a general belief in the degree to which people and institutions (i.e., schools) can be considered reliable and supportive. Without adequate levels of trust and positive attitudes, Chu et al. hypothesized that students have little incentive to engage in effortful performances on standardized tests. Chu et al. contended that a student's fear of potential failure could create protective self-regulatory behavior that could negatively impact intellectual risk taking in relation to standardized tests. Chu et al. also acknowledged the connection between trust and effort that students put forth as identified in expectancy theory. This theory postulates that individuals (students in this instance) who trust their leaders (teachers in this instance) are more likely to engage in positive behaviors that will lead to desired outcomes. The desired outcome would be improved performance on standardized assessments that measure mathematical success. Chu et al. cautioned that teachers and researchers might be cautious in applying too much weight to large scale standardized test results and the validity of test scores might be highly influenced by the attitudes of the test takers. Regardless of how much mathematical knowledge and skill students learn, Chu et al. concluded that they may or may not possess the trust and positive attitude necessary to invest significant energy and effort in test taking. Perhaps students in an examination-focused culture like Singapore have a different approach to test taking than students from a less examination-focused culture like the United States. This study is important because it addresses the possible impact of incentives such as serious skills-promoting games on the engagement of students in standardized testing.

In a significant study about the relationship of female teachers' mathematics anxiety to mathematics achievement for female students, Beilock et al. (2010) postulated that female students might inherit mathematical anxiety from female teachers in early elementary school. Beilock et al. determined the level of mathematical anxiety for grade one and grade two teachers by examining their responses to the *Mathematics Anxiety Rating Scale*. Beilock et al. also measured mathematics achievement for female students who were enrolled in these teachers' classes by analyzing their responses to the applied problems section of the *Woodcock Johnson III Tests of Achievement*. In addition, Beilock et al. assessed the gender beliefs of these female students by analyzing the drawings they created after listening to gender-neutral stories. Beilock et al. found that female students are more likely to suffer from mathematics anxiety than male students. Beilock et al. also found that at the beginning of the year, female students' mathematics achievement did not correlate with the level of anxiety their teachers expressed about mathematics. However, at the end of the year, female students' mathematical achievement directly correlated to the anxiety of their teachers whereas male students' scores did not. This study is important to gender-based research about mathematics because educators need to consider the idea that female students often conform to teacher stereotypes that they may not be as good at mathematics as male students.

In a related study, Hembree (1990) conducted a meta-analysis of 151 research studies about the nature, effects, and relief of mathematics anxiety of students in Grades 1-12 and at the postsecondary level. Hembree found that students in Grades 5 through 12 with high anxiety about mathematics demonstrated low motivation for learning mathematics. In addition, Hembree found that high anxiety correlated with low levels of concept mastery and learning mathematics in general. Hembree noted that for in students in Grades 5 through 12, the relationship between anxiety and aptitude was stronger for male students than female students but the relationships were weaker at the postsecondary level. Hembree also noted that junior high and high school students with high levels of mathematics anxiety enrolled in fewer high school and college mathematics classes. In addition, Hembree found male high school students with high levels of mathematics anxiety showed more avoidance behaviors than female students with high levels of mathematics anxiety and were less likely to enroll in mathematics classes at the postsecondary level. However, female students displayed higher levels of mathematics anxiety than male students from grade six to the postsecondary level. Hembree also examined mathematics anxiety by college classes (though not by gender) and found that students with the highest levels of anxiety were enrolled in four classes, including mathematics for elementary teachers, developmental mathematics, elementary education, and remedial algebra. Preservice elementary school teachers demonstrated the highest levels of mathematics anxiety. This study is important because it shows how anxiety correlates with learning mathematics and that female students may exhibit higher anxiety levels than male students. In addition, preservice elementary school teachers may exhibit mathematical anxiety that could negatively affect their students.

In summary, gender disparities in STEM career fields may continue to persist, even if more females are recruited to enroll in college level STEM courses (Grunspan et al., 2016). Because female students often exhibit more mathematical anxiety than male students (Beilock et al., 2010; Schommer-Aikens et al., 2015) and because high mathematical anxiety correlates with low motivation and learning (Beilock et al., 2010; Hembree, 1990; Schommer-Aikens et al., 2015), mathematical anxiety might be one type of psychological stress contributing to gender inequities in mathematics courses related to the STEM pipeline. Another possible contributing factor to gender inequities might be the persistence of gender stereotyping (Alcock et al., 2014; Casad et al., 2015; Plante et al., 2009). A third possible reason for persistent gender inequities in mathematics could be student attitudes and perspectives towards mathematics (Chu et al., 2014; Grunspan et al., 2016; Johnson et al., 2012; Mendick, 2005; Pahlke et al., 2014). However, even though Grunspan et al. (2016) and Casad et al. (2015) acknowledged the negative effects of gender stereotyping on student learning, Pahlke et al. (2014) was uncertain about the degree of negative impact gender stereotyping has on students. Stoet and Geary (2012) also concluded that not enough data supports stereotype threat as the primary cause of gender differences in students' mathematics success. Lindberg et al. (2010) also concluded that gender alone could not account for gender differences in students' mathematics performance. From a different perspective, Alcock et al. (2014) found individual personalities had a greater predictive effect on students' mathematical achievement than gender. In a similar vein, Johnson et al. (2012) found that male and female students react differently to stereotype threat and stereotype lift. Picho et al. (2013) also examined stereotype threat and concluded that it exerts a negative effect on the mathematics testing performance of female students and the negative effect was stronger for female students in middle and high school than in college. Plante et al. (2009) questioned whether or not male and female students believe mathematics to be in the male or female domain and called for more research in that area. Schommer-Aikens et al. (2015) found that student beliefs serve as adequate predictors of mathematical performance and mathematical anxiety. Similarly, Guo et al. (2015) noted that improving student motivation results in improved student mathematics achievement, regardless of gender or socioeconomic status. Similarly, Gaspard et al. (2015) found that female students reported mathematics learning as more useful for education than male students, but less useful for future personal goals than male students. Gaspard et al. (2015) suggested that teachers close the gender gap in mathematics performance by emphasizing the importance of mathematics in daily life and in various career fields. Spitzer and Aronson (2015) found that female students lose confidence in STEM subjects as they get older. If gender discrepancies emerge in the data for this study, gender stereotyping and its concomitant gender stereotype threat may not be the direct cause (Casad et al., 2015; Stoet & Geary, 2012). However, there may be no significant gender inequities revealed in my analysis of the data. Research from Pahlke et al. (2014) and Lindberg et al. (2010) raises that possibility.

Gender Differences on Large Scale Mathematical Assessments

Mathematical achievement can be demonstrated through formative and summative assessments (Adabor, 2013; Doig, 2006; Schoenfeld, 2015). Formative and summative assessments have "complementary objectives of determining mathematics learning outcomes" (Adabor, 2013, p. 56). Formative assessments are given to students before or during instruction in order to determine student strengths and areas of improvement (Adabor, 2013; Doig, 2006). The primary purpose of formative assessments is to provide both students and their teachers with feedback as to areas that need improvement before instruction is completed and a summative assessment is given (Schoenfeld, 2015). Teachers may use data from formative assessments to "build up student understanding through focusing on student thinking while engaging in rich mathematical tasks" (Schoenfeld, 2015, p. 183). In contrast, students complete summative assessments at the end of instruction (Adabor, 2013; Schoenfeld, 2015). The primary purpose of summative assessments is to assess student knowledge after instruction by assigning scores to individual students that are compared to other student scores (Schoenfeld, 2015). Summative assessments may be external, taking the form of national, state, or district-level assessments (Adabor, 2013) or may be internal, such as end-of-course examinations (Schoenfeld, 2015). Adabor (2013) suggested that teachers and students view formative assessments as "micro-summative" assessments because they assess student knowledge after smaller amounts of instructional time than a full course (p. 56). In relation to summative assessments, standardized testing, whether from nationally-normed instruments, such as the NAEP, SAT, and Terra Nova, or from international comparisons, such as the PISA and the TIMSS, measure student performance in mathematics that may be different from classroom performance, which is often measured by formative assessments. Large-scale assessments provide national and international data about mathematical performance and gender gaps. Researchers use data from these large-scale assessments to examine both short-term and long-term trends in gender equity or inequity in order to inform educational policymakers as to the effectiveness of current policy. Researchers also review large-scale assessments for patterns of gender equity in mathematics. Even when test scores reflect significantly different results for male students and female students, causative factors other than gender alone may be accountable. Because this study is focused on an examination of numbers of students scoring above the national average on TNTE scores, which are large-scale summative assessment scores, this section includes only an analysis of current research on large-scale summative assessments for international and national assessments.

International Assessments

Mathematical assessments may improve student learning by providing teachers and other stakeholders with student performance data to improve instruction. Data derived from national and international assessments are useful when comparing student performance among countries, and their results may contribute to understanding perceived deficiencies in mathematics curricula and instructional practices. The TIMSS assessment has been recognized as an appropriate measure for international comparisons of mathematical achievement for students in Grades 4, 8, and 11 (Kaleli-Yilmaz & Hanci, 2015; Wagemaker, 2002).

In an examination of variables related to the TIMSS mathematics results, Kaleli-Yilmaz and Hanci (2015) determined that gender was a neutral element in mathematical performance for grade eight students. Kaleli-Yilmaz and Hanci found gender differences among countries in relation to TIMSS scores. Male students scored higher than female students on this assessment in some countries, and female students scored higher than male students on this assessment in other countries. Kaleli-Yilmaz and Hanci found similar results when examining gender in the three cognitive domains that the TIMSS measures, which includes knowing, applying, and reasoning. They found that student performance varied by country. For example, from the 2011 TIMSS assessment, students in Turkey demonstrated the highest level of performance in the reasoning domain and the lowest level of performance in the knowing domain. On the other hand, students in the United States demonstrated stable scores. In the knowing domain, male and female students in the United States scored the same (519) and nearly the same in the applying domain (500 for female students and 506 for male students) and in the reasoning domain (501 for female students and 506 for male students). Students in Ghana, on the other hand, scored lowest in all three domains for both female and male students. Kaleli-Yilmaz and Hanci found that female students scored higher than male students in the cognitive domain. They also found that even though female students demonstrated higher levels of performance in all three cognitive domains than male students, the differences were not significant. Overall, Kaleli-Yilmaz and Hanci found that student mathematical performance did not change relative to gender. In fact, the variable more responsible for high student performance on the TIMSS assessment in Turkey was the level of education attained by the parents. Kaleli-Yilmaz and Hanci also found no alignment between the Turkish curricular materials and the TIMSS assessment questions. Furthermore, they noted that educators in countries with above average TIMSS scores were more likely to use constructivist instructional practices while Turkish teachers used intense rote learning practices. Kaleli-Yilmaz and Hanci recommended that Turkish educators adopt a constructivist approach to improve mathematics performance for their students. This research is significant because constructivism is a popular approach to mathematics instruction that may be gender blind.

In other related research about international assessments in mathematics, Cheema and Galluzzo (2013) explored the gender gap in mathematics achievement for K-12 public school students in the United States. Cheema and Galluzzo used multiple regression analysis to examine data from the PISA for an American population of 4,733 students in Grades 4, 8, and 12. They found a small but significant gender gap in mathematics, even when they controlled for various demographic characteristics, such as socioeconomic status and race. However, this gender gap persisted when they controlled for female students’ anxiety levels and self-efficacy. This research is particularly important because it affirms the idea that gender differences exist in student mathematics performance. This research also brings the issue of gender bias into the discussion about STEM education that should be addressed at the elementary and secondary school levels.

In another study about international assessments in mathematics, Stoet and Geary (2013) analyzed gender differences from a decade of PISA mathematics and reading tests that involved approximately 1.5 million 15-year-old students from 75 countries. The tests were administered in 2000, 2003, 2006, and 2009. Stoet and Geary discovered that in the majority of the 75 countries involved in the study, male students continued to score higher than female students in mathematics. Stoet and Geary also found that female students typically scored higher than male students in reading and lower than male students in mathematics both across nations and within nations. Stoet and Geary also discovered a consistent and strong inverse relationship between the size of the gender gap in mathematics and the size of the gender gap in reading. The countries that reported larger gender gaps in mathematics also reported smaller gender gaps in reading. However, the countries that reported smaller gender gaps in mathematics reported larger gender gaps in reading. Although Stoet and Geary discovered this relationship, they were unable to ascertain why this correlation occurred. Notably, Stoet and Geary found that the gender gap was the largest at the highest end of the continuum (e.g. the students scoring at the 95th percentile and above) with male students consistently outperforming female students. Stoet and Geary suggested that the gender gap in the STEM fields could be related to the fact that students entering STEM fields often score in the highest percentiles on mathematics tests such as the PISA. In addition, although the number of female students performing in the highest percentiles on the mathematics portion of the PISA increased in the last 2 decades, the gender gap related to students scoring at the top quartile has remained stable. This study is important because it demonstrates a trend of gender inequity in mathematics performance on the PISA over the last decade, especially in the top quartile. This study is also important because an inverse relationship was found between gender gaps in reading and mathematics, particularly in relation to the performance of top quartile students. This finding is particularly notable in relation to the gender gap in STEM because in the 1980s, the ratio of male students to female students in secondary school mathematics was 13:1 (Benbow & Stanley, 1980, 1983). By the 1990s, this ratio had dropped to 4:1, but it has not decreased since (Wai, Lubinski, Benbow, & Steiger, 2010). Students who score in the top quartile on the mathematics portion of the PISA are often students who enter STEM fields and more of those students are men than women (Stoet and Geary, 2013).

Using the 1999 TIMSS data, Birenbaum et al. (2005) analyzed the mathematics performance of grade eight students in the United States, Israel, and Singapore. Birenbaum et al. found that the mathematics performance of grade eight students in Singapore was superior to the mathematics performance of grade eight students in both the United States and Israel. Even so, because of dissatisfaction with the lack of creativity demonstrated by their graduates, the Singapore government began to reconstruct their education system to foster creativity, autonomy, and flexibility. This research is important because it illuminates how a national government shared responsibility in improving student outcomes in mathematics.

National Assessments

In a significant study, Reilly et al. (2014) conducted a meta-analysis of NAEP assessment data from 1990 to 2011, specifically focusing on gender gaps in science and mathematics achievement in the United States. Their work was predicated on the assumption that to raise the underrepresentation of female students in STEM fields, concerted and sustained efforts are required. After referencing previous research confirming no significant gender differences in general intelligence, Reilly et al. reiterated well-documented gender differences in quantitative reasoning related to science and mathematics skills. Acknowledging that multiple theories exist to explain gender differences in reasoning (i.e., biological, psychological, social, and psychobiological theories), Reilly et al. aligned their meta-analysis more closely with Hyde's (2005) gender similarities hypothesis, arguing that more similarities than differences exist in the cognitive abilities of the sexes. Reilly et al. based their analysis of the NAEP assessment data on four key questions: (a) Does a gender gap exist in science and math achievement and, if so, is the gap diminishing? (b) Do male students show more variability in performance? (c) If sex differences exist in means and in variance, how do they combine to affect the proportion of male and female students attaining advanced proficiency in math and science? and (d) If there are differences between the genders, are they present in earth, physical, and life science achievement? Reilly et al. noted that the NAEP is administered to students in Grades 4, 8, and 12 on rotating cycles of 2 to 3 years for mathematics and four to five years for science. For this meta-analysis, the sample included 1,925,100 students for the mathematics assessment and 878,916 students for the science assessment. Reilly et al. chose a random-effects model over a fixed-effects model to analyze the assessment data because it provided wider confidence levels and better estimates about how much variability was present across the data samples. Reilly et al. found statistically significant gender differences in quantitative reasoning at all three grade levels. Additionally, Reilly et al. found a moderate overrepresentation of high achieving male students in Grades 4 and 8, and a considerable increase in differences for high achieving male students in Grade 12. A similar disparity was found in the physical science and earth science assessments, but not for the biological science assessments. Reilly et al. concluded that their analyses supported the research literature that reported a widening gender gap for middle school students. They also noted, however, that the National Center for Education Statistics conducted a NAEP-TIMSS comparison and found the assessment frameworks comparable, which does not imply that the measurements of gender performance in mathematics are comparable. Because their conclusions confirmed gender differences exist in mathematical achievement, Reilly et al. recommended educational interventions in relation to achieving gender equity in mathematics and science achievement. This research is particularly relevant because mathematics and science are integral components of STEM and because they confirmed gender differences in mathematical achievement.

In other related research, Lindberg et al. (2010) performed a meta-analysis of the research literature to examine trends in gender and mathematics performance. They examined data from 242 studies that represented a sample of over one-million people from preschool to postsecondary and large data sets from the past 20 years including the *National Longitudinal Surveys of Youth,* the *National Education Longitudinal Study of 1988,* the *Longitudinal Study of American Youth,* and the NAEP. Lindberg et al.'s goal was to answer the following six questions: (a) What is the magnitude of gender differences? (b) What is the magnitude of test effects relative to the depth of knowledge of the testers? (c) At what age do gender differences appear or disappear? (d) What is the magnitude of variations relative to ethnic groups? (e) Has the magnitude of gender differences declined since 1990? and (f) Do male adults and young men display greater variance in scores? Lindberg et al. found no overall gender differences in mathematics performance on these tests. They also discovered that female students and male students performed equally well on mathematics tests when they examined student performance for all ages. Their analysis of the second part of their study found similar results, and they drew the conclusion that gender cannot be used as a reliable predictor of mathematics performance. Lindberg et al. also concluded that data from their meta-analysis do not support the argument that gender-integrated classrooms put female students at a disadvantage in terms of learning. However, they noted that at the high school level, a gender gap favors male students in relation to the number of students who scored at the top quartile on the NAEP. Lindberg et al. suggested that other variables might account for gender bias and lingering gender stereotypes. This study is important because it presents strong data illustrating that cultural shifts have occurred in the United States that may have erased gender gaps in current mathematics performance. However, in the top quartile, gender bias favoring males at the high school level may still exist.

In an earlier but significant study of large-scale mathematics assessment, Doig (2006) described assessment practices in several English-speaking countries, including the uses of summative and formative assessments. Doig noted an inherent problem with the reporting of large-scale assessments is that teachers may be reluctant to understand the results as formative in nature. Scale anchoring (i.e., attaching meaning to a scale identifying what students seemingly know and can do) is used by the NAEP and the TIMSS, and developmental continua are used in Australia and the United Kingdom to describe and monitor student progress in particular academic domains. Both scale anchoring and developmental continua are attempts to utilize data from large-scale summative assessments to impact formative assessments and instructional improvements in the classroom. In referencing earlier research, Doig (2006) acknowledged that although large-scale and system-wide assessments are focused at a macro-level and are summative in nature, they also provide micro-level data on sub-groups. For example, micro-level data on the gender or the non-English speaking backgrounds of students can be determined from these assessments. In addition, Doig maintained that the results of large-scale assessments such as the TIMSS, the NAEP, the PISA, and the annually administered Key Stage 2 and 3 Mathematics Test for England and Wales, although summative in nature, could be used to improve classroom instructional practices. This study is important because the results of large-scale assessments in mathematics could be used to improve classroom instruction, which in turn could increase student achievement.

In summary, the key findings related to research on international and national assessments in mathematics are that there are gender differences in mathematical success at these levels (Cheema & Galluzzo, 2013; Kaleli-Yilmaz & Hanci, 2015; Stoet & Geary, 2013). In addition, the gender gap is largest, favoring male students, at the highest level of mathematics achievement (Stoet & Geary, 2013). These studies are important because they present strong data illustrating that cultural shifts have occurred relative to mathematics instructional practices and student performances within recent decades. Collectively, these studies support the notion of using large-scale assessments to reveal or track shifts in mathematics performance by gender.

Mathematics Achievement of Military Children

A lack of current research exists on mathematics achievement, particularly in relation to gender, for military-connected students. However, there is some research comparing DoDEA schools, which have high populations of military-connected students, with other public schools or with state averages (Abell, 2004; National Center for Education Statistics, 2015). DoDEA schools provide support systems to reduce the impact of these stressors on student achievement (Department of Defense Education Activity, 2015). Teachers in DoDEA schools are trained to be aware of the unique needs and stressors that children of military families face and to help them overcome those stressors (Department of Defense Education Activity, 2014b). In addition, a crisis management team (CMT) comprised of trained professionals is present in every DoDEA school to help DoDEA students and teachers in the event of a crisis (Abell, 2004). One purpose of the CMT is to provide support before, during, and after each deployment (Abell, 2004). The members of the CMT also understand the variety of stressors that military children face (Abell, 2004).

In relation to mathematics achievement, military children, whether stateside or overseas, routinely experience unique stressors unlike children in non-military families (Angrist & Johnson, 2000; Bradshaw et al., 2010; Burrell, Adams, Durand, & Castro, 2006; De Pedro et al., 2011; Engel, Gallagher, & Lyle, 2010; Esqueda et al., 2012).

These stressors may impact their learning in mathematics. Concerning the measurement of mathematics achievement for military-connected children, current research examining NAEP and Terra Nova mathematics scores are important to this study because they are nationally-normed tests. In addition, because DoDEA students take the Terra Nova and NAEP assessments, researchers can compare DoDEA student scores with those scores of other non-military students who took the assessments at the same time but were not exposed to the same stressors as military children as well as with non-DoDEA military-connected students who were in public school systems. Therefore, in this section, current research is analyzed in relation to the unique stressors that military children face and the mathematics achievement of military-connected students in relation to Terra Nova and the NAEP.

Unique stressors. Researchers have indicated that the psychological health, behavior, and academic achievement of children of military service members are often negatively impacted by frequent moves, frequent parental deployments, feelings of insecurity due to uncertainty about the future, and worries about parents being injured or killed in war zones (Angrist & Johnson, 2000; Bradshaw et al., 2010; Burrell, Adams, Durand, & Castro, 2006; De Pedro et al., 2011; Engel, Gallagher, & Lyle, 2010; Esqueda et al., 2012; Lester et al., 2016). Educators in schools with few military-connected children may not be aware of the unique stressors that these students face, and therefore, they may not respond appropriately to their needs (Esqueda et al., 2012). In contrast, military children who attend overseas schools and who face the same stressors as their counterparts in stateside schools tend to function well academically (Beardsley, 2015; De

Pedro et al., 2011; Department of Defense Education Activity, 2010; Department of Defense Education Activity, 2011a; Esqueda et al., 2012). This may be due to unique circumstances within the overseas communities. Researchers have found military-connected students who have strong and healthy social connections through family as well as support from both formal systems and informal networks had less depression, more persistence, better self-regulation, and increased academic performance (Lucier-Greer, Arnold, Mancini, Ford, & Bryant, 2015).

Children with one or more parents in the military face unique emotional stressors such as frequent parental absence due to deployments, being uprooted from extended family and friends because a parent has received orders to move to a new base, anxiety about parents deployed to war zones, and cultural challenges such as adapting to living in foreign countries (Creech & Hadley, 2014; Department of Defense Education Activity, 2014a; Lester et al., 2016; Lucier-Greer et al., 2015; Richardson, Mallette, O'Neal, & Mancini, 2016). The stress of deployment, especially to hostile war zones, can lead to parental depression, PTSD, traumatic brain injury (TBI), and divorce (Lemmon & Stafford, 2014; Lester et al., 2016; Lucier-Greer et al., 2015; Negrusa, Negrusa, & Hosek, 2014). Mothers, in addition to fathers, can be deployed (Negrusa et al., 2014); in fact, approximately 30% of active duty women are mothers (Lester & Flake, 2013). In a study about the effects of work-related absences on military families during the Gulf War, Angrist and Johnson (2000) found a statistically significant correlation between an increase in female soldiers being deployed to the Gulf War and an increase in divorces. Similarly, Negrusa et al., (2014) found that female service members deployed to hostile areas for extended lengths of time had a higher divorce rate than their male counterparts. In one of the first studies examining the impact of wartime parental deployments and subsequent reintegration on families with young children's social and emotional adjustment, Lester et al. (2016) found parents who had posttraumatic stress and depression due to deployments or due to the deployments of their spouses were more likely to have children with high anxiety levels, social emotional problems, and academic problems. Lester et al. found that the additional stressors placed on civilian spouse during and after a dangerous deployment could traumatize the spouses, which could mean children could have both parents suffering from depression or PTSD symptoms. In addition, Lester et al. found a correlation between increased deployments to warzones and instabilities in marriages as well as impaired family functioning. Children of enlisted members were at an increased risk for difficulties than children of officers. In 2013, approximately 52,322 military members were in marriages with military spouses while an additional 155,000 military members raised children as single parents (Lester & Flake, 2013). Children of dual military deployed parents or single parents who are deployed are often placed with temporary caregivers who are friends or relatives (Creech & Hadley, 2014). Children who may already be worried about their parents often face the additional stress of learning new house rules, getting along with other children in their new family, and learning new routines. If a parent is deployed to a war zone, children face an additional stressor of understanding that their parent could be wounded or killed (Barker & Berry, 2009). Children of parents who are or have been deployed may exhibit a variety of negative behavior and socioemotional problems, including depression, anxiety, and sleep disorders (Barker & Berry, 2009).

In a meta-analysis of current research about the impact of military deployment and reintegration on parenting, Creech and Hadley (2014) found five current studies linking parental deployment and increases in child abuse. Creech and Hadley found an increase in binge drinking for middle school and high school adolescents living with temporary caregivers due to parental deployment compared to civilian students. They also found that secondary students with parents who are deployed are more likely to abuse alcohol, marijuana, and prescription drugs than nonmilitary counterparts. Similarly, in an earlier study about parental deployment and adverse effects on children, Gibbs, Martin, Kupper, and Johnson (2007) found a correlation between parental deployment and increased child abuse or neglect, which was greater if the deployment was combat-related.

In contrast involving more recent research, Card et al. (2011) conducted a meta-analysis of 16 studies published from 1978 to 2010 relating to military deployment and military children's adjustment involving 19,172 participants; 12 of those studies included 2,707 participants who reported emotional symptoms such as anxiety and depression. Card et al. found a lack of methodologically rigorous research related to internalizing or externalizing behaviors of military children with deployed parents. They also found only small correlations between deployment and overall behavior or emotional problems in military children with the largest effect in middle childhood. In addition, Card et al. were only able to locate five studies that demonstrated correlations between deployment and negative academic outcomes, but when they attempted to complete a qualitative synthesis of those five studies they were unable to determine the magnitude of that association. In addition, Card et al. were only able to locate two studies that used standardized assessments as a measure of academic achievement of students with deployed parents, both of which found small but statistically significant negative associations. This study is relevant because it demonstrates a gap in the research on military children and academic achievement while simultaneously revealing a lack of rigorous studies relating to deployment and children's behavioral and emotional problems that could lead to academic problems.

Children of military parents may also be uprooted from their homes and schools because their parents are reassigned to new locations. De Pedro et al. (2011) explored district, school, and community perspectives of the experiences of military students and found that school-aged military children are subject to an average of between 6 to 9 moves between kindergarten and high school. In fact, students of military families have a 31% mobility rate every year (Department of Defense Education Activity, 2015).

In an earlier study about the unique stressors that military children face, Lyle (2006) examined the effects of parental military deployments and military household relocations on military children's standardized mathematics test scores. The sample size included mathematics scores from the *Texas Assessment of Academic Skills* for 13,000 students in Grades 3 through 8. All participating students had at least one military parent in the United States Army who was stationed in Texas in 1997 or 1998. Lyle found that both male and female students with a mother in the United States Army scored lower in mathematics than students with a father in the United States Army. Additionally, Lyle compared the mathematics performance of children of enlisted soldiers to the children of officers on the *Texas Assessment of Academic Skills*. Lyle discovered that children of officers demonstrated a smaller standard deviation in mathematics scores and scored 5-6 standard score points higher in mathematics than the children of enlisted soldiers. Lyle also found male students scored slightly lower than female students in mathematics if they were children of enlisted soldiers, but no significant difference was found between the male and female children of officers. This study is important because gender differences were found in the mathematics scores for children of enlisted soldiers but not for children of officers.

Terra Nova results. In a study about military deployments and students' academic achievement, Engle et al. (2010) examined academic achievement on the Terra Nova for 56,116 students in Grades 3 through 11 enrolled in DoDEA schools between 2002 and 2005. All students in this study had at least one parent who was enlisted in the United States Army during the time of the study. Engle et al. directly linked parental military deployment to significant reductions on Terra Nova test scores in mathematics, science, reading, social studies, and English language arts. For example, they found that students who had a deployed parent scored 0.42% lower than students who did not have a parent deployed. Engle et al. did not find a statistical difference between mathematics scores on the Terra Nova for students in DoDEA stateside schools compared to students in DoDEA overseas schools. Similarly, Engle et al. did not find a statistical difference on mathematics Terra Nova scores of students with one military parent compared to students in dual military families. Engle et al. also examined the effects of the length of deployments on educational outcomes. They found an 8-month parental deployment corresponded with a decline of students' mathematics scores of 1.50% and a decline of the total score of 0.90%. However, the average length of deployment at the time of the study was almost double at 15 months. A 15-month deployment corresponded with a -1.68% total NCE score decrease and a drop in the mathematics NCE score of -2.82%. They also found a 0.92% reduction in mathematics scores if the parent was deployed during the actual month of testing. In fact, even if a deployed parent had returned, students with parents who were absent closer to the testing date scored lower on mathematics NCE scores than students who had parents return earlier in the year. Perhaps more worrisome, Engle et al. found that the adverse effects of parental deployment on Terra Nova scores could persist for several years, and in some cases, did not completely disappear for 4 to 5 years. Engle et al. concluded that modest detrimental effects in all subjects were found if students had parents deployed during the year of testing and tended to dissipate after the parents' return; however, minor detrimental effects could persist for several years. They also concluded that the statistically significant effects of parental deployment on Terra Nova scores were more pronounced if the parent was deployed during the month of testing or had participated in a lengthy deployment, and they found that mathematics scores were lower than total scores. Engle et al. speculated that the adverse effects of deployment on the test scores of military children in non-DoDEA schools might be more severe than on the test scores of military children in DoDEA schools because educators in DoDEA schools might be better equipped to manage deployments and the stressors that accompany these deployments. Engle et al. concluded that the negative effects of parental deployment on military children's academic learning could have implications for educational policy and could even affect national security. They suggested that educators in schools with significant populations of military children should create programs that alleviate these adverse effects of deployment on their academic achievement. This study is important because the findings indicate that parental absence due to military deployments may have a negative effect on the academic performance of military children.

NAEP results. Assessment of student performance in mathematics in public and private schools, including DoDEA schools, at the national level has been limited (National Center for Education Statistics, 2015). As the only standardized test administered at the national level to K-12 students, the NAEP is a defining assessment (National Center for Education Statistics, 2015). The NAEP is not administered every year and is only administered to students in Grades 4, 8, and 12 (National Center for Education Statistics, 2015). For example, the mathematics portion was administered to students in Grade 12 in 2013 and to students in Grades 4 and 8 in 2015 (National Center for Education Statistics, 2015). In addition, results for private school students were included in the 2013 results, but not in the 2015 results, because the private schools who chose to participate did not meet the minimum number of students that the NAEP participation guidelines required (National Center for Education Statistics, 2015). The results are reported and compared by state as well as by subcategories, such as race, gender, and disability in a report titled *The Nation's Report Card* (National Center for

Education Statistics, 2015). The sample grade four population from 2015 included approximately 279,000 students, and the sample grade eight population included approximately 273,000 students (National Center for Education Statistics, 2015). In 2015, the NAEP scores for both grades were higher than the original 1990s assessments but slightly lower than the previous scores recorded in 2013 (National Center for Education Statistics, 2015).

In an analysis of current NAEP data, mathematics scores for female students in Grades 4 and 8 were lower in 2015 than in 2013 (National Center for Education Statistics, 2015). Female students scored lower in mathematics than male students in Grades 4, 8, and 12 in 2013. In spite of the unique challenges military-connected students may face, military-connected students in DoDEA schools consistently score above the national average in mathematics on the NAEP (Abell, 2004; National Center for Education Statistics, 2015). In 2015, female students in DoDEA schools in Grades 4 and 8 scored higher in mathematics than their non-DoDEA counterparts. However, at both grade levels, the percentage of DoDEA female students who scored at the advanced level in mathematics paralleled the national scores. DoDEA female students also scored higher than non-DoDEA female students in mathematics in 2015 (National Center for Education Statistics, 2015). However, DoDEA female students scored lower in mathematics than male students enrolled in public and private schools and male students enrolled in DoDEA schools in 2015 (National Center for Education Statistics, 2015).

The NAEP recognizes mathematical proficiency and mastery of complex and challenging mathematical subject matter when students correctly answer 250 out of 500 questions on the grade four assessment and correctly answer 300 out of 500 questions on the grade eight assessment (National Center for Education Statistics, 2015). NAEP results for 2015 indicated that only 40% of students in Grade 4 achieved proficiency or above in mathematics (National Center for Education Statistics, 2015). By Grade 8, the number of proficient students dropped to 33%, and only 26% of Grade 12 students achieved proficiency (National Center for Education Statistics, 2015). Scores for female students have risen from 213 points in 1990 when the NAEP was first administered to 241 points in 2013 before dropping to 239 points in 2015; however, for the 25-year time span that the NAEP has been administered, average scores for both female students and male students in Grades 4 and 8 have never reached a score of proficient or above (National Center for Education Statistics, 2015).

In relation to gender, NAEP results for grade four male students closely parallel their female counterparts, although the scores in mathematics for male students are slightly higher (National Center for Education Statistics, 2015). For instance, in 1990, female students averaged 213 points while male students averaged 214 points (National Center for Education Statistics, 2015). The highest average scores in mathematics were reported in 2013 with an average of 241 points for female students and 242 points for male students (National Center for Education Statistics, 2015). In 2015, scores for both male and female students dropped slightly, but male students still retained a slight lead with 241 points compared to 239 points for female students (National Center for Education Statistics, 2015). The mathematics scores of female students in grade eight also improved over the 25-year time period. Female students averaged 262 points in

1990 (male students averaged 263 points), and they earned the highest score of 284 points in 2013 (male students averaged 285 points). In 2015, female students in grade eight tied with male students at 282 points, still below the required score of 300 or above to be considered proficient in mathematics (National Center for Education Statistics, 2015).

NAEP scores in mathematics from 1996 to 2015 included DoDEA schools (National Center for Education Statistics, 2015). In 1996, 21% of students nationally were proficient in mathematics, and 19% of DoDEA students were proficient in mathematics. By 2013, 45% of DoDEA students were proficient in mathematics, which was higher than the national average of 42%. By 2015, 49% of DoDEA students were proficient in mathematics compared to the national average of 40%.

These NAEP results also indicated that a higher percentage of grade four students achieved proficient or above scores in mathematics than grade eight students. For DoDEA schools, male students scored higher than female students in mathematics at both grade levels. At grade four, 42% of male students scored at or above proficient compared to 38% of female students. In addition, 9% of male students and female students at both of these grade levels scored at the advanced level in mathematics. At grade eight, 34% of male students scored at or above proficient in mathematics compared to 33% of female students, indicating that the gender gap is closing. However, only 6% of grade four female students (compared to 10% of males) and 8% of grade eight female students (compared to 9% of males) scored at the advanced level in mathematics, indicating the gap is largest at the advanced level of achievement with male students scoring higher than female students (National Center for Education Statistics, 2015).

NAEP scores also indicated that a small gender gap in mathematics scores exists at the national level with 38% of grade four female students (compared to 42% of male students) and 33% of grade eight female students (compared to 34% of male students) scoring at or above proficient (National Center for Education Statistics, 2015). In relation to DoDEA schools, 47% of grade four female students and 39% of grade eight female students scored at or above proficient in mathematics. In comparison, 51% of grade four male students and 40% of grade eight male students scored at or above proficient in mathematics. A gender gap also exists among the students who scored at the highest level in mathematics. Only 10% of grade four male students enrolled in DoDEA schools achieved the advanced level in mathematics, compared to 9% of grade four male students nationally while grade eight male students enrolled in DoDEA schools tied the national average of 9%. However, a smaller percentage of female students at both grade levels reached the advanced level in mathematics with grade four female students enrolled in DoDEA schools scoring at 6% nationally and at 3% for female students enrolled in DoDEA schools. For grade eight female students at the national level, 33% (1% less than the male students) scored at or above the proficient level and 8% (1% less than the male students) scored at the advanced level. In contrast, 39% (6% higher than the national average) of the DoDEA grade eight female students scored at or above proficient and 8% scored at the advanced level. Female DoDEA students in Grades 4 and

8 scored slightly below DoDEA male students, but above female students in relation to the national average (National Center for Education Statistics, 2015).

In summary, military children face a variety of unique stressors that could negatively impact their mathematical achievement (Angrist & Johnson, 2000; Bradshaw et al., 2010). Several studies have been conducted on the impact of stressors such as frequent mobility and deployment of parents on military children's wellbeing and/or academic achievement (Angrist & Johnson, 2000; Bradshaw et al., 2010; Burrell et al., 2006; De Pedro et al., 2011; Engel et al., 2010, Esqueda et al., 2012). Despite the unique situations that military families face, DoDEA students consistently perform above the national average on standardized mathematics tests such as the NAEP and Terra Nova (National Center for Education Statistics, 2015). NAEP scores also indicate that female DoDEA students in Grades 4 and 8 score slightly below DoDEA male students in Grades 4 and 8 on the NAEP in relation to mathematics, indicating that female students may be more responsive to the unique stressors of military life than male students (National Center for Education Statistics, 2015). However, both female students and male students in the DoDEA school system score above the national average on standardized math assessments but not to the same degree (National Center for Education Statistics, 2015).

Summary and Conclusions

In summary, this chapter included a review of the literature, beginning with a description of the literature search strategies that I used to conduct this review. In addition, a detailed description was included of the theoretical framework, which was based on Sax's (2005, 2009, 2010, 2011) meta-analysis of research concerning genetic, anatomical, and functional differences between male students and female students and how those differences impact their learning. Sax contended that male students and female students demonstrate differences in learning styles, vision, hearing, behavior at play, and even attitudes toward homework and that teachers should address these differences in their instruction. The review of literature was divided into three sections. The first section, gender differences in mathematical learning, included an analysis of current research about student perspectives and attitudes about mathematics and mathematical abilities, gender stereotyping, and psychological stressors. The second section included an analysis of current research about gender differences in large scale international and national assessments. The third section, mathematics achievement of military-connected children, included an analysis of current research about the unique stressors that military children face, mathematics achievement in schools as defined by the Terra Nova results, and mathematics achievement in schools as defined by the NAEP results.

Three major themes emerged from this review of the literature. The first major theme was that gender differences have been demonstrated in mathematical learning, even though these differences continue to be contentious. Some researchers maintain that male students and female students experience mathematical anxiety differently (Beilock et al., 2010; Hembree, 1990; Hill et al., 2016; Schommer-Aikens et al., 2015). Other researchers contend that male students and female students experience stereotyping in relation to gender and concomitant psychological stress differently (Alcock et al., 2014; Casad et al., 2015; Pahlke et al., 2014; Picho et al., 2013; Plante et al., 2009). These differences might contribute to gender inequities in mathematics achievement, which may result in gender gaps in the STEM fields. Other researchers in this area contend that gender stereotyping alone cannot definitively account for differences in mathematical performance (Casad et al., 2015; Grunspan et al., 2016; Guo et al., 2015; Lindberg et al., 2010; Stoet & Geary, 2012). Other researchers also maintain that a plausible reason for continued gender inequities in mathematics could be differences in student perspectives and attitudes towards mathematics (Chu et al., 2014; Grunspan et al., 2016; Johnson et al., 2012; Mendick, 2005; Pahlke et al., 2014). For example, Guo et al. (2015) found that improved student motivation, regardless of gender, improved mathematics scores for all students and Alcock et al. (2014) found that student personalities were a greater predictor of mathematical achievement than gender. Similarly, Spitzer and Aronson (2015) discovered that female students lose confidence in their mathematical ability as they get older, and Gaspard et al. (2015) found that female students believe mathematics is less useful for future career goals than male students.

A second theme that emerged from this literature review was that gender differences have been found in relation to large-scale assessment scores in mathematics. Research at international and national levels revealed that gender differences in mathematical achievement exist, but they vary by country (Reilly et al., 2014; Stoet & Geary, 2013). A review of the research on large-scale assessments at international and national levels also affirmed the hypothesis that gender differences in student performance in mathematics are persistent, though the gender gap may be diminishing and/or gender may be a neutral causative factor (Kaleli-Yilmaz & Hanci, 2015; Lindberg et al., 2010; Reilly et al., 2014; Stoet & Geary, 2013). Research also shows an inverse relationship exists between reading and mathematics scores with international averages showing that female students tend to score higher than male students in reading (Stoet & Geary, 2013). However, male students continue to significantly outnumber female students in relation to mathematics scores on assessments conducted in the United States, particularly in the highest percentiles (e.g. 95th percentile and above) for high school students on the PISA and NAEP (Kaleli-Yilmaz & Hanci, 2015; Lindberg et al., 2010; Stoet & Geary, 2013). Analyses of large scale assessments also showed that gender differences in mathematical achievement have resulted in recommendations that countries at the national level (e.g. Turkey, Singapore) modify or adjust their mathematics instruction in order to decrease the gender gap. For example, Kaleli-Yilmaz and Hanci (2015) found that teachers in countries who use the constructivist approach produce above average mathematics scores on the TIMSS, implying that constructivism may be a desirable instructional approach in achieving gender equity in mathematical teaching and learning. In an examination of mathematics achievement on the PISA, Cheema and Galluzo (2013) found only a small but significant gender gap in mathematics and expressed optimism for addressing this gap at the elementary and secondary school levels. Even though large-scale assessments indicated that Singapore students in grade eight outscored American and Israeli students in grade eight in mathematics, the Singapore government sought to modify their educational program to further improvements in other areas, demonstrating how the national government can take responsibility for student achievement (Birenbaum et al., 2005). However, conflicting interpretations of large-scale assessment data still persist. Reilly et al. (2014) noted that although no significant differences exist in general intelligence between genders, the gender gap in mathematics at the middle school level may be widening. On the other hand, Lindberg et al. (2010) conducted a meta-analysis of data from over one million people from preschool to postsecondary in relation to large data sets from the past 20 years including the *National Longitudinal Surveys of Youth,* the *National Education Longitudinal Study of 1988,* the *Longitudinal Study of American Youth,* and the NAEP. They found negligible gender differences in mathematics performance when examining all ages. However, at the highest percentiles for high school students, Lindberg et al. found a slight gender difference with more male students scoring in the top percentiles than female students. At the postsecondary level, they found slightly more female than male students scored in the top percentiles. Notably, Lindberg et al. also found that gender-integrated classrooms did not put female students at a disadvantage in learning mathematics.

A third theme in the literature was that educators in DoDEA schools are often successful in meeting the mathematical learning needs of military-connected children because DoDEA students consistently outperform the national average on standardized tests of mathematics achievement. Whether educated in DoDEA schools or non-DoDEA schools, children in military families are faced with unique stressors (e.g. frequent moves, parental deployments, feelings of insecurity and worries about parental well-being in war zones) that can account for negative psychological, behavioral, and academic outcomes (Angrist & Johnson, 2000; Bradshaw et al., 2010; Burrell et al., 2006; De Pedro et al, 2011; Engel et al., 2010; Esqueda et al., 2012). Frequent parental deployments, often six to nine times in a student's K-12 school experience, resulted in an annual mobility rate of 31% for the military-connected student population (Department of Defense Education Activity, 2014a). Family moves often cause military children to be estranged from their deployed parent(s) and extended family members and friends, and they are required to adjust in order to live in a foreign country with additional cultural challenges (Creech & Hadley, 2014; Department of Defense Education Activity, 2014a). Within the military community, differences in mathematics achievement have also been found between children of officers and children of enlisted parents (Lyle, 2006). Yet DoDEA students consistently scored above the national average on the NAEP in mathematics (Abell, 2004; National Center for Education Statistics, 2015). Engle et al. (2010) also found significant reductions in Terra Nova mathematics scores for DoDEA students when one or more parent was deployed. *The Nation's Report Card* also confirmed a drop in DoDEA mathematics scores when parents were deployed (National Center for Education Statistics, 2015). DoDEA students were first included in the NAEP testing in 1996 when 21% of all students in the United States were proficient in mathematics and 19% of these students attained that level (National Center for Education Statistics, 2015). Since then, DoDEA students have continually improved and surpassed the national average in mathematics (National Center for Education Statistics, 2015). While female DoDEA students score higher than the national average in mathematics, they have consistently scored slightly below male DoDEA students. The gap is closing but remains largest at the advanced level (National Center for Education Statistics, 2015). DoDEA students also function well academically compared to military students in non-DoDEA schools and consistently score above the national average in mathematics on national assessments such as the Terra Nova and NAEP (Beardsley, 2015; De Pedro et al., 2011; Department of Defense Education Activity, 2010, 2011b; Esqueda et al., 2012; National Center for Education Statistics, 2015). The 2015 NAEP mathematics scores for students in Grades 4 and 8 placed DoDEA students at the top of the nation, scoring in second place (tied with seven states) for Grade 4 and in fourth place for Grade 8 (O'Gara & Kanellis, 2015). These strong achievement results for DoDEA students may be related to the support systems that DoDEA educators have established (Department of Defense Education Activity, 2014b). For example, teachers in the DoDEA school system receive specialized professional development which may help students cope with their students' unique stressors (Department of Defense Education Activity, 2014b). Additionally, the DoDEA school system provides crises management teams to assist both students and teachers in times of crises (Abell, 2004).

A major gap found in the research literature was a lack of longitudinal data on the mathematics achievement of high achieving military-connected students, particularly in relation to gender equity. Another gap found in the literature was a lack of research focusing on the impact of parental deployment on the performance of male students and female students in mathematics. A third gap was a lack of research on mathematics achievement for military-connected children in public schools because many public-school systems do not identify or monitor military children in their student populations (De Pedro et al., 2011). A final gap revealed a lack of data on the educational outcomes of military-connected children in general (Esqueda et al., 2012). My results from the present study contribute to filling the gaps in the literature related to longitudinal data on the mathematics achievement of high achieving military-connected students in relation to gender equity.

In chapter 3 I included a description of the research method that I used to conduct this quantitative study. In this chapter, I discussed the research design and rationale, sampling and sampling procedures, and use of archival data. I also described the data analysis plan, threats to validity, and ethical procedures.

Chapter 3: Research Method

The purpose of this quantitative study was to describe the relationship between student gender and mathematics achievement on the TNTE for military-connected students in Grades 3 through 9 who attended schools located in the participating school district from 2012 to 2016. Therefore, this chapter includes a description of the research method, including the research design and the rationale for choosing that design. Sampling procedures and procedures for archival data collection are also described as well as the data analysis plan. Threats to validity and ethical procedures are also discussed.

Research Design and Rationale

In order to answer the research questions and hypotheses for this quantitative study, a correlational research design was selected. In correlational studies, quantitative researchers examine the potential relationship between two or more variables to determine if there is a relationship or covariation (Waters, n.d.). A correlational research design was particularly suited for this study because the archived data were numerical, and these data were collected using a measurable instrument and were statistically analyzed using logistic regression in addition to an ANOVA and a two-tailed *t* test. Furthermore, correlational research has been effectively used in several studies relating to mathematics (Chafin et al., 2015; Tatar, Zengin, & Kağızmanli, 2015; Yenilmez & Turgut, 2016; Zorrilla-Silvestre, Presentación-Herrero, & Gil-Gómez, 2016). Therefore, this research design was selected because it was the best design to answer the research questions and hypotheses for this study.

Independent variables for this study were gender and grade band, and the dependent variable was the percentage of military-connected students in Grades 3 through 9 who were enrolled in the participating school district during the years 2012 to 2016 and scored in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE. Location was not chosen as a variable because the student population was mobile, and, therefore, students might be enrolled in a number of schools during their educational experiences. Because de-identified anonymous archival data were used, there was not a way to track individual students, and, therefore, using location as a variable might have resulted in unreliable data analysis. This correlational research design choice was consistent with other quantitative research designs that researchers use to advance knowledge in the field because the goal of the researcher was to determine if there was a relationship between the independent variables of gender and grade band and the dependent variable of percentage of military-connected students in Grades 3 through 9 who were enrolled in the participating school district during the years 2012 to 2016 and scored in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE. In addition, it was an appropriate design because the data came from instrument-based assessments, and the results were analyzed and interpreted using statistical analyses.

The target population for this study included all high achieving military-connected students in Grades 3 through 9 who attended schools in the participating school district and who completed the TNTE mathematics test during the years 2012 to 2016. However, the exact number of schools and student populations were different for different years in this study because some of the schools consolidated and student populations fluctuated. This change meant individual students and schools could not be tracked for comparison within the scope of this study. There were approximately 195 schools and 135,571 students in this study.

Sampling and Sampling Procedures

The sampling strategy that was used to conduct this study was a convenience sample that included total mathematics scores from high achieving military-connected students in Grades 3 through 9 who were enrolled in schools in the participating school district and completed the TNTE mathematics assessment during the years 2012 to 2016. High achieving for the purposes of this study was defined as scoring at or above the national average in mathematics on the TNTE.

In relation to inclusion criteria, only TNTE NCE mathematics test scores for high achieving military-connected students in Grades 3 through 9 who were enrolled in schools in the participating district during the years 2012 to 2016 were included in this study. The years 2012 to 2016 were chosen because they provided a robust amount of current data. The grade bands were chosen to represent elementary and post-elementary levels.

Concerning sampling procedures, the number of students in the top two quartiles (at or above the national average score of 50) were examined. A large sample size was selected because it was more likely to result in statistically significant data (Frankfort-Nachmias & Nachmias, 2008). However, a large sample size could have produced statistically significant results even if the differences were small. According to

Frankfort-Nachmias and Nachmias (2008), researchers need to choose a large enough sample size so that the data can be used to generate empirically supported generalizations. To accurately estimate the unknown parameters from known statistics, the sample population must be defined, the sample size must be determined, and the sample design must be explained (Frankfort-Nachmias & Nachmias, 2008). The sample population for this study included high achieving military-connected students in Grades 3 through 9 who were enrolled in schools in the participating school district and who took the mathematics subtests for the TNTE in 2012 to 2016.

The four main types of probability samples are (a) a simple random sample, (b) a systematic sample, (c) a stratified sample, and (d) a cluster sample (Frankfort-Nachmias & Nachmias, 2008). None of these four probability sample types were selected because the goal was to obtain a stronger picture of the entire population than any of these sample types allow. By using the entire population, sampling errors were negated, and population estimates were increased. In addition, because archival data were selected for this study, a convenience sample was a good choice for a sampling procedure. As Frankfort-Nachmias and Nachmias (2008) noted, "Ideally, the sampling frame should include all the sampling units in the given population [but] such information is rarely available" (p. 165). Although it may not always be possible or convenient for researchers to include an entire population within a study, because archival data were used for this study, a summary of each grade level for the years 2012 to 2016 (2012-2015 for Grade 9) was available as a convenience sample. In addition, although the population consisted of students from approximately 195 schools, only four numbers (the total number of students in each of the top two quartiles by gender) per grade level per year were used in the analysis, resulting in 20 numbers per grade level except for Grade 9 because ninth grade students did not take the assessment in 2016 (136 total numbers), which was a manageable amount of data.

Summary data from all 135,571 students was used. I analyzed the summary data from year to year and indicated the total number of students in each of the top two quartiles by gender. Because the data were aggregated, I used the percentages for statistical analysis, and there were four numbers analyzed for each grade level: two female numbers (percentages at and above the standard) and two male numbers (percentages at and above the standard). For example, for the year 2012, there were 1,198 third grade females who scored at the standard, representing 24.57% of the total third grade population within my study for the year 2012. Similarly, there were 1,248 females (25.60%) who scored above the standard, 1,150 males who scored at the standard (23.58 %), and 1,280 males who scored above the standard (26.25%). The four percentages (25.60, 23.58, 23.58, and 26.25) were the four scores used for the year 2012 for Grade 3. For purposes of statistical analysis of data, the percentages were used because the total numbers of students fluctuated from year to year and grade to grade.

An online a-priori sample size statistics calculator version 4.0 created by Soper (2017) was used to conduct a power analysis to determine appropriate sample size. An anticipated small effect size (Cohen's *d*) of 0.2 was originally chosen, along with a probability level of 0.05, and a desired statistical power level of 0.8. For a two-tailed hypothesis, a minimum of 788 students needed to be included in the total sample size. In addition, a minimum of 394 students needed to be included in each group. In order to obtain a range of numbers, a power analysis was run for an anticipated medium effect size of 0.5. The probability level of 0.05 and desired statistical power level of 0.8 remained the same. For an anticipated medium effect size and a two-tailed test, the minimum number of students needed would be 128 with a minimum of 64 students in a group. Because the final data provided by the participating school district included 136 groups, the medium effect size was chosen.

Procedures for Archival Data Collection

Due to time restraints and access to data, I chose to request de-identified archival data, which represented a large sample size over a 5-year period. The archival data that were used was the number of students who scored in the top two quartiles from a standardized mathematical test called the TNTE. This test was administered to military-connected students in Grades 3 through 9 who were enrolled in schools in the participating district and who completed this test between 2012 and 2016. The TNTE represented the best source of archival data because it was a standardized test that was nationally normed, and the majority of the military-connected student population in the participating school district completed the test during the selected years. Therefore, a large amount of data was available. In addition, by using archival data for a secondary analysis, researcher bias was reduced, and the reliability and validity of this study were increased. In addition, requesting de-identified data was the best ethical choice because the participating school district's Research Center protected the personal information and anonymity of all schools and students. Therefore, readers of the study will also not be able to connect scores with specific students or schools.

I sought approval to conduct this study from both the Institutional Review Board (IRB) at Walden University and from the Research and Evaluation Board from the participating school district. In order to acquire access to the archival data, I first obtained conditional approval from the participating school district's Research Center to collect data. After I received conditional approval from the participating school district's Research Center in the form of a signed letter of cooperation indicating the agreement of the participating school district to provide data, I obtained approval from the Walden University IRB. After I had approval from the Walden IRB, I sent a copy of the approval letter back to the participating school district's Research Center and requested access to the data. The participating school district's Research Center provided me with a letter of cooperation but requested that identifying information be removed before publication of this dissertation. Therefore, the letter of cooperation was removed from Appendix A and replaced with a note stating the original letter of cooperation is on file with Walden University.

In relation to the mathematics test, CTB/McGraw-Hill developed the TNTE. The TNTE mathematics tests that are administered to students in grades K to 12 are considered valid and reliable because they were normed in 2007 in a nationwide, empirical study (CTB/McGraw-Hill, 2008). The pilot study involved over 57,000 students, and the standardization sample involved over 275,000 American students (Brown & Coughlin, 2007). Predictive validity for the tests administered to students in

Grades 3 through 11 was established via a study that linked it to Pennsylvania State System of Assessment scores (Brown & Coughlin, 2007). In addition, content and construct validity were established through comprehensive state curriculum guides, examination of national standards, and examination of current textbooks (Brown & Coughlin, 2007). The reliability coefficient values for internal consistency for the TNTE are 80% to 95%, which shows strong internal validity, and the coefficient value for the standard error of measurement was 2.8 to 4.5 and showed variability in the standard error of measurement across grade levels. However, standard errors were approximately 0.25 to 0.33 standard deviation units, which is small (Brown & Coughlin, 2007). A permission letter to use TNTE from CTB/McGraw-Hill was not included in the appendix because I used archival data to examine the numbers of students who scored in each quartile, and, therefore, I was not administering the TNTE or examining individual test scores. The use of TNTE mathematics scores was justified and appropriate for this study because it is a valid and reliable nationally normed and standardized test and because the participating school district's Research Center collected results on this test during the 5-year period involved in this study.

Data Analysis Plan

A logistic regression analysis, a two-tailed *t* test, and an ANOVA were conducted to answer the following research questions and related hypotheses as stated in Chapter 1.

Research Question (RQ) 1: Is there a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for normal curve equivalent (NCE) mathematics scores on the TNTE using gender as a predictor?

$H_0 1$: There is no statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE using gender as a predictor.

$H_a 1$: There is a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE using gender as a predictor.

RQ2: Is there a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE by gender and grade band?

$H_0 2$: There is no statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE by gender and grade band.

$H_a 2$: There is a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE by gender and grade band.

De-identified student scores were released as a dataset from the participating school district's Research Center. I used SPSS version 24.0 software (IBM Corp., 2016) to conduct statistical analyses. I used logistical regression using gender as a predictor and I used logistical regression to examine gender and grade band. I also used ANOVA and a two-tailed *t* test to compare the mean scores for the number of male students and the number of female students in each quartile on the mathematics portion of the TNTE. The power was set at 0.8 and the Type I error rate at 0.05.

The dataset had already been generated as part of the participating school district's organizational operations. The data were cleaned and screened by the participating school district's Data Center before the data were released to me, according to their research protocols. The participating school district's Data Center de-identified the data and presented me with only the total numbers of male and female students who scored in each quartile at each grade level for each year. No student names or schools were attached to the data. I examined the number of female and male students each year who scored at each of the two top quartiles (levels) by grade level (3 through 9) to determine if there were any statistically significant differences between the genders by grade level and quartile.

Threats to Validity

TNTE mathematics scores were chosen as an existing measure of mathematical progress because validating a new instrument was beyond the scope of this study. TNTE was standardized and nationally normed to increase reliability and validity. I reduced researcher bias and threats to validity by requesting de-identified archival data and by using a large existing dataset that covered a 5-year timespan. The statistical tests that I used to conduct this study were logistical regression, ANOVA and a two-tailed *t* test.

External Validity

Creswell (2009) identified the following three threats to external validity for quantitative research studies: interaction of selection and treatment, interaction of setting and treatment, and interaction of history and treatment. The threat of interaction of selection and treatment means that due to the "narrow characteristics of participants in the experiment, the researcher cannot generalize to individuals who do not have the characteristics of the participants" (Creswell, 2009, p. 165). To reduce this threat, I restricted "claims about groups to which the results cannot be generalized" (Creswell, 2009, p. 165). The threat of interaction of setting and treatment means that due to specific characteristics of the research setting, results might not be generalizable to participants in different settings. I reduced this threat by including all of the participating school district's schools. Finally, the threat of interaction of history and treatment means that "because results of an experiment are time-bound, a researcher cannot generalize the results to past or future" (Creswell, 2009, p. 165). This threat can be reduced by replicating the study at a different time (Creswell, 2009, p. 165). I reduced this threat by including five years of data.

Internal Validity

The ten most common threats to internal validity include history, maturation, regression, selection, mortality, diffusion of treatment, compensatory/resentful demoralization, compensatory rivalry, testing, and instrumentation (Creswell, 2009). The threat of history was reduced because all student groups completed a grade-level appropriate TNTE test annually during a 5-year period. The threat of maturation was reduced because it was assumed that the participating school district's students received similar curriculum as they progressed through the grades. The threat of regression was reduced because "scores over time regress to the mean" (Creswell, 2009, p. 163). The threat of regression could be reduced further if outliers were removed. It was assumed that outliers were not removed because the numbers that were analyzed were the total numbers of the participating school district's students who scored in the top two quartiles. Therefore, because only total numbers were analyzed, outliers could not be eliminated, which would be located in the last quartile. The threat of selection was also reduced because all mathematics test scores were included and because students were randomly assigned to schools and may have moved during the study. Creswell (2009) noted that "characteristics have the probability of being equally distributed among the experimental groups" (p. 163). The threat of mortality was reduced because a large sample size was selected, which should "account for dropouts" (Creswell, 2009, p. 163). The threats of diffusion of treatment, compensatory/resentful demoralization and compensatory rivalry were not applicable threats to this study. The threat of testing was also not assumed applicable to this study because students were administered a different test for each grade level, and therefore, they could not become familiar with test answers. The threat of instrumentation was reduced because the same test (the third edition) was administered to students at each grade level annually.

Construct Validity

Construct validity is a way to define the ability of an experiment to actually measure what it claims to measure (Shuttleworth, 2009). Construct validity is similar to external validity (Shuttleworth, 2009). However, if an experiment has strong construct validity, Shuttleworth (2009) contended that the experiment addresses the variable(s) being tested. Creswell (2009) noted that construct validity threats occur when researchers fail to adequately define and measure variables. Construct validity concerns were reduced because I clearly defined and measured the variables.

Ethical Procedures

Three important ethical issues are the obtaining, treatment, and storage of data. I obtained approval from the Walden University IRB and also from the participating school district's Research Center before obtaining data. I chose secondary data analysis using de-identified archival data to eliminate bias. By asking for identifiers to be removed before I collected the data, I reduced risk to the students and their military-connected parents. The data were already generated as part of normal operations for the participating school district. I requested de-identified anonymous data, which helped to eliminate bias and protect the schools and their students from any accidental release of their personal information. I did not know which scores came from which schools.

In terms of data storage, I will keep all data electronically and in paper format. I will store paper copies in a locked file cabinet, and I will destroy this data after five years of the publication of this study. I will store electronic copies on a password encrypted laptop and delete them after five years, as requested by Walden University.

Another ethical concern was that I was an employee of a school in the participating district, and I understood that conducting a study in my own work environment could pose a potential conflict of interest. I did not accept any monetary or other incentives for this research. I chose to further reduce the risk of conflict of interest by using anonymous archival data that the participating school district's Research Center provided and by using a correlational research design to reduce any potential bias I might have had. I did not conduct this study during my duty day, and I did not anticipate any positive or negative outcomes related to the publication of this research and my teaching job. I also did not provide any incentives because I used archival data.

Summary

The purpose of this chapter was to describe the methodology that was used to collect and analyze data in relation to the purpose of this study, which was to explore the relationship between gender and mathematics achievement for high achieving military-connected students in Grades 3 through 9 who were enrolled in the participating school district's schools during the years 2012 to 2016. In this chapter, the correlational research design and the rationale for that design was described. This chapter also included a description of the sampling and sampling procedures and the procedures for collecting the archival data needed for this study. In addition, the data analysis plan was described as well as threats to external and internal validity and to construct validity. Specific strategies to enhance validity included using a nationally normed testing instrument, collecting data from multiple schools, using 5 years of data, and using de-identified archived data to reduce researcher bias. Finally, ethical procedures were discussed, including using de-identified data, data handling, and data storage.

In chapter 4 I included a discussion of the results of this study. The chapter begins with a description of how the archival data were collected by educators in the participating school district's schools and the process used to obtain this data. In addition, a detailed analysis of the archival data was included. Key findings of the study were discussed in relation to the research question and the hypotheses. Finally, a discussion of the reliability and validity of this study was included.

Chapter 4: Results

Introduction

The purpose of this study was to describe the relationship between gender and mathematics achievement on the TNTE for high achieving military-connected students in Grades 3 through 9 who were enrolled in schools in the participating school district from 2012 to 2016. All of the students in the study scored in the top two quartiles or above, meaning they scored at the national average of 50 per cent or above on the TNTE. An explanation of how the individual grade levels were examined is addressed in Chapter 4. The quantitative research design chosen for this study was correlational, using historical data from the participating school district's Research Center. The research questions and hypotheses investigated in this study were the following:

Research Question (RQ) 1: Is there a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for normal curve equivalent (NCE) mathematics scores on the TNTE using gender as a predictor?

$H_0 1$: There is no statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE using gender as a predictor.

$H_a 1$: There is a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE using gender as a predictor.

RQ2: Is there a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE by gender and grade band?

$H_0$2: There is no statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE by gender and grade band.

H_a2: There is a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE by gender and grade band.

Chapter 4 consists of a description of data collection including demographics and how the data were collected. Next, the results section includes a description of the statistical tests conducted. I also discuss why I used the ANOVA to take into consideration the difference between males and females and conducted some tests by gender. In addition, I discuss why more advanced tests were not conducted. Finally, a summary of the results answering the research questions is provided. I conclude with a transition statement to Chapter 5.

Data Collection

I used archival third-party data collected by the participating school district from 2012 to 2016. The data were provided by the participating school district's Research Center. Data collected was a convenience sample consisting of the number of military-connected students (by gender and grade) who scored in the top two quartiles for NCE mathematics scores on the TNTE within the participating school district from approximately 195 schools. The data were de-identified prior to releasing them for this study. The sample was representative of the population of interest because it consisted of all 135,571 of the students in Grades 3 through 9 who were military-connected and scored in the top two quartiles for NCE mathematics scores on the TNTE from 2012 to 2016.

Upon approval from the participating school district's Research Center and approval from Walden University's IRB #02-05-18-0187505, IBM's SPSS Statistics Version 24 (IBM Corp., 2016) was used to conduct all data analysis. Gender was coded 1 for female and 2 for male. Grade was coded 1 through 7, with 1 representing third grade and 7 representing ninth grade. Years were coded 1 through 5 in order, with 1 representing 2012, the first year of data, 2 representing 2013, and so on. The standards were coded 1 for the third quartile and 2 for the fourth (and highest) quartile. Percent was entered numerically to two decimal places.

As noted in Chapter 3, an online a-priori sample size statistics calculator version 4.0 created by Soper (2017) was used to conduct a power analysis to determine appropriate sample size. An anticipated small effect size (Cohen's *d*) of 0.2 was chosen, along with a probability level of 0.05, and a desired statistical power level of 0.8. For a two-tailed hypothesis, a minimum of 788 students need to be included in the total sample size. In addition, a minimum of 394 students need to be included in each group. In order to obtain a range of numbers, a power analysis was also run for an anticipated medium effect size of 0.5. The probability level of 0.05 and desired statistical power level of 0.8 remained the same. For an anticipated medium effect size and a two-tailed test, the minimum number of students needed was 128 with a minimum of 64 students in a group. I originally anticipated being able to analyze the data for a small effect size because the data were from over 100,000 students. After collection, data were categorized into 136 (68 male, 68 female) aggregated representative grades, which represented the percentages of the summaries of all 135, 571 students scoring at or above the standard. Therefore, I chose to examine a medium effect size. The data collected consisted of the numbers of students who scored at or above the standard on the TNTE mathematics assessment. In total, there were 135,571 students broken down as follows: 23,969 third graders, 21,213 fourth graders, 20,676 fifth graders, 19,283 sixth graders, 19,312 seventh graders, 18,646 eighth graders, and 12,472 ninth graders. In Chapter 3, I noted anticipating 140 total representative grades representing 20 representative grades per grade level. However, after I obtained the data, I learned that the ninth graders did not take the assessment in 2016, so there were only 4 years of data for ninth grade, resulting in 136 representative grades to analyze: 60 representative grades for grades 3 through 8 and eight representative grades for grade 9 for each gender, totaling 136 representative grades which were analyzed. The numbers analyzed were the total percentages of females or males at each grade level for every year. For example, in 2012, there were 1,198 female students in grade 3 who scored at the standard, representing 24.57 per cent of the total number of grade 3 students who took the test and scored at or above the standard. The number 24.57 was used as one representative grade to represent all 1,198 female grade 3 students who scored at the standard for year 2012. Following this protocol, there were 10 representative grades for female students in grade 3: one for each year 2012 to 2016 representing the percentage of female students scoring at the standard and one for each year representing the percentage of female students scoring above the standard (from the total number of students scoring at or above the standard) on the TNTE. Similarly, there were 10 representative grades for male students in each grade level from grades 3 through 8. There were only 8 representative grades for students in grade 9 because they did not take the TNTE in 2016. The 136 representative grades used for statistical analysis are listed in Tables 1 and 2 in the columns labeled per cent. Table 1 shows the total numbers and percentages of female and male elementary students (Grades 3 through 5) by grade and year who scored at the standard or above the standard. The percentages were the representative grades analyzed in my statistical analysis. Table 2 shows the same data for female and male post-elementary students (Grades 6 through 9).

Table 1

Numbers of Elementary Students

	Numbers of elementary female students				Numbers of elementary male students			
Year	At the standard	%	Above the standard	%	At the standard	%	Above the standard	%
<u>Grade 3</u>								
2012	1198	24.57	1248	25.60	1150	23.58	1280	26.25
2013	1145	23.55	1262	25.96	1058	21.76	1397	28.73
2014	1177	24.17	1242	25.51	1051	21.59	1399	28.73
2015	1107	22.68	1319	27.02	1021	20.91	1435	29.39
2016	1046	23.35	1149	25.65	1000	22.32	1285	28.68
Total	5673	23.67	6220	25.95	5280	22.03	6796	28.35
<u>Grade 4</u>								
2012	1103	25.24	1086	24.85	1016	23.25	1165	26.66
2013	1079	24.75	1143	26.22	1013	23.24	1124	25.79
2014	983	23.51	1115	26.66	861	20.59	1223	29.24
2015	857	20.94	1196	29.22	825	20.16	1215	29.68
2016	871	21.15	1154	28.02	821	19.93	1273	30.91
Total	4893	23.16	5694	26.96	4536	21.47	6000	28.41
<u>Grade 5</u>								
2012	1044	23.77	1146	26.09	967	22.02	1235	28.12
2013	997	22.84	1183	27.10	996	22.82	1189	27.24
2014	1015	24.18	1162	27.68	884	21.06	1137	27.08
2015	900	22.66	1090	27.44	834	21.00	1148	28.90
2016	886	23.63	1016	27.10	831	22.17	1016	27.10
Total	4842	23.42	5597	27.07	4512	21.82	5725	27.69

Table 2

Numbers of Post-elementary Students

	Numbers of post-elementary female students				Numbers of post-elementary male students			
Year	At the standard	%	Above the standard	%	At the standard	%	Above the standard	%
Grade 6								
2012	920	23.37	999	25.37	969	24.61	1049	26.64
2013	971	24.58	977	24.73	931	23.56	1072	27.13
2014	971	24.31	981	24.56	903	22.60	1140	28.54
2015	879	23.35	1051	27.92	781	20.74	1054	27.99
2016	824	22.67	903	24.84	830	22.83	1078	29.66
Total	4565	23.67	4911	25.47	4414	22.89	5393	27.97
Grade 7								
2012	983	24.13	1014	24.90	967	23.74	1109	27.23
2013	894	22.62	1035	26.18	934	23.63	1090	27.57
2014	920	24.17	992	26.06	857	22.52	1037	27.25
2015	879	22.80	1028	26.67	812	21.06	1136	29.47
2016	845	23.31	1017	28.06	742	20.47	1021	28.17
Total	4521	23.41	5086	26.34	4312	22.33	5393	27.93
Grade 8								
2012	960	23.94	1018	25.39	891	22.22	1141	28.45
2013	892	23.24	993	25.87	854	22.25	1100	28.65
2014	813	22.10	1002	27.24	797	21.67	1066	28.98
2015	766	21.80	980	27.89	717	20.40	1051	29.91
2016	774	21.47	1018	28.24	716	19.86	1097	30.43
Total	4205	22.55	5011	26.87	3975	21.32	5455	29.26
Grade 9								
2012	709	21.72	859	26.32	638	19.55	1058	32.41
2013	660	20.30	946	29.10	631	19.41	1014	31.20
2014	623	20.61	845	27.95	582	19.25	973	32.19
2015	582	19.84	831	28.32	565	19.26	956	32.58
2016	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Total	2574	20.64	3481	27.91	2416	19.37	4001	32.08

For purposes of statistical analysis of data, the percentages were used because the total numbers of students were not equal from year to year or grade to grade. As stated earlier, there were 10 representative grades for females and 10 representative grades for males for each grade from 3-8 and 8 representative grades for each gender for grade 9, resulting in a total of 68 representative grades for each gender and 136 representative grades total. The numbers in the columns labeled percent in Tables 1 and 2 above add up to the 136 representative grades used for statistical analysis.

Results

In order to answer the first research question, I conducted an independent sample *t* test examining all of the data from all of the grade levels and years (representing all 135,571 student scores aggregated into 136 representative grades representing the percentages for each grade level per year as indicated in Tables 1 and 2) was conducted using SPSS Version 24 using percent as the test variable and gender as the group variable (IBM Corp., 2016). Table 3 shows the group statistics for the *t* test.

Table 3

Group Statistics for t test

	Gender	*N*	Mean	Standard deviation	Standard error mean
Percent	Female	68	24.8095	2.26552	.27473
	Male	68	25.1908	3.90993	.47415

The confidence interval percentage for the *t* test was 95%, meaning the alpha value was .05. The results were, $t = .696$, $df = 134$, $p = .000$. Because $.000 < .05$, the variances are not assumed to be equal. The *Sig* (*p*-value) for the *t* test was .488. Because

.488 > .05, I could accept the null hypothesis. Therefore, I concluded that there was no statistically significant evidence to support my first research question and accepted the null hypothesis that there was no statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE using gender as a predictor.

In order to answer my second research question, I conducted an ANOVA and a logistical regression. Students were analyzed by gender for the ANOVA to see if there was a difference in scores by grade level for females or males. First, I ran an ANOVA to analyze all of the female students with grade level as a fixed factor. Tables 4 through 7 show the results of the tests of between subject effects, estimates of grade levels, and univariate tests for female students and Tables 8 through 11 show the results for male students. There were 68 representative grades per gender: 10 representative grades (two per year representing the percentage of students by gender scoring at and the percentage of students by gender scoring above the standard compared to the total number of students scoring at or above the standard) for each grade from 3-8 and eight representative grades for Grade 9 from the years 2012-2016. There were only eight representative grades per grade level for Grade 9 because the students in Grade 9 were not administered the assessment in 2016. It is important to note that the 136 representative grades analyzed represent all 135,571 student scores. Table 4 shows the results of the between-subjects effects tests. The dependent variable was the representative grades, meaning the percentage of female students out of the total number of students who scored at or above the standard by grade level. It is thought that the population means of the different grade levels are not all equal (i.e., at least one is different from the others). This can be tested using multiple two-sample t-tests to compare all the pairs. But if each test is 0.05, the probability of making a Type 1 error when running three tests would increase.

A better method is ANOVA (analysis of variance), which is a statistical technique for determining the existence of differences among several population. I ran an ANOVA between-subjects effects test because it is an extension of the between-groups *t* test to examine two or more groups simultaneously. The ANOVA compares two types of variances: the variance within each sample and the variance between different samples. The between-subjects factors were the grade levels. Students were analyzed by gender to see if there was a difference between the grade levels and within the gender groups. It was determined that grade level was not a determining factor for scores when between-subjects test was conducted, $F(.168)$, $p = .984$, as seen in Table 4.

Table 4

ANOVA Test of Between-Subjects Effects for Female Students

Source	Type III Sum of squares	*df*	Mean square	*F*	*Sig.*
Corrected model	5.590[a]	6	.932	.168	.984
Intercept	41548.540	1	41548.540	7491.933	.000
Grade	5.590	6	.932	.168	.984
Error	338.292	61	5.546		
Total	42198.799	68			
Corrected total	343.882	67			

Note. [a] R Squared = .016 (Adjusted R Squared = -.081)

Next, I ran an estimated marginal means test for female students. The mean was 24.794 with a standard error of .286. The estimated marginal means test was run to reconfirm there were no outliers. There were no outliers.

Then I ran the estimates for grade levels. Table 5 shows the results of the estimates for grade levels. The dependent variable was the percent of female students scoring at or above the standard. As indicated on page 118, there were 68 female representative grades used: two per year per grade level. The first representative grade used was the percentage of females scoring at the standard in relation to the total number of students scoring at or above the standard. The second representative grade used was the percentage of females scoring above the standard in relation to the total number of students scoring at or above the standard.

Table 5

ANOVA Estimates for Grade Levels for Female Students

			95% Confidence interval	
Grade	Mean	Std. error	Lower bound	Upper bound
Grade 3	24.806	.745	23.317	26.295
Grade 4	25.056	.745	23.567	26.545
Grade 5	25.249	.745	23.760	26.738
Grade 6	24.570	.745	23.081	26.059
Grade 7	24.890	.745	23.401	26.379
Grade 8	24.718	.745	23.229	26.207
Grade 9	24.270	.833	22.605	25.935

As shown in Table 5 above, I ran the estimates for grade levels test to see if there was a difference between the grade levels. There was not a difference between grade levels. Next, I ran the pairwise comparisons test. The dependent variable was labeled percent, and was the representative grade representing the percent of female students who scored at or above the national standard compared to the total number of students in the data set scoring at or above the national standard. The pairwise test was run to examine individual grades. There was no significant difference between individual grades. Table 6 shows the results of the pairwise test.

Table 6

ANOVA Pairwise Comparisons for Female Students

(I) Grade	(J) Grade	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
Third Grade	Fourth Grade	-.250	1.053	1.000	-3.589	3.089
	Fifth Grade	-.443	1.053	1.000	-3.782	2.896
	Sixth Grade	.236	1.053	1.000	-3.103	3.575
	Seventh Grade	-.084	1.053	1.000	-3.423	3.255
	Eighth Grade	.088	1.053	1.000	-3.251	3.427
	Ninth Grade	.536	1.117	1.000	-3.006	4.078
Fourth Grade	Third Grade	.250	1.053	1.000	-3.089	3.589
	Fifth Grade	-.193	1.053	1.000	-3.532	3.146
	Sixth Grade	.486	1.053	1.000	-2.853	3.825
	Seventh Grade	.166	1.053	1.000	-3.173	3.505
	Eighth Grade	.338	1.053	1.000	-3.001	3.677
	Ninth Grade	.786	1.117	1.000	-2.756	4.328
Fifth Grade	Third Grade	.443	1.053	1.000	-2.896	3.782
	Fourth Grade	.193	1.053	1.000	-3.146	3.532
	Sixth Grade	.679	1.053	1.000	-2.660	4.018
	Seventh Grade	.359	1.053	1.000	-2.980	3.698
	Eighth Grade	.531	1.053	1.000	-2.808	3.870
	Ninth Grade	.979	1.117	1.000	-2.563	4.521
Sixth Grade	Third Grade	-.236	1.053	1.000	-3.575	3.103
	Fourth Grade	-.486	1.053	1.000	-3.825	2.853
	Fifth Grade	-.679	1.053	1.000	-4.018	2.660
	Seventh Grade	-.320	1.053	1.000	-3.659	3.019
	Eighth Grade	-.148	1.053	1.000	-3.487	3.191
	Ninth Grade	.300	1.117	1.000	-3.242	3.842
Seventh Grade	Third Grade	.084	1.053	1.000	-3.255	3.423
	Fourth Grade	-.166	1.053	1.000	-3.505	3.173
	Fifth Grade	-.359	1.053	1.000	-3.698	2.980
	Sixth Grade	.320	1.053	1.000	-3.019	3.659
	Eighth Grade	.172	1.053	1.000	-3.167	3.511
	Ninth Grade	.620	1.117	1.000	-2.922	4.162
Eighth Grade	Third Grade	-.088	1.053	1.000	-3.427	3.251
	Fourth Grade	-.338	1.053	1.000	-3.677	3.001
	Fifth Grade	-.531	1.053	1.000	-3.870	2.808
	Sixth Grade	.148	1.053	1.000	-3.191	3.487
	Seventh Grade	-.172	1.053	1.000	-3.511	3.167
	Ninth Grade	.448	1.117	1.000	-3.094	3.990
Ninth Grade	Third Grade	-.536	1.117	1.000	-4.078	3.006
	Fourth Grade	-.786	1.117	1.000	-4.328	2.756
	Fifth Grade	-.979	1.117	1.000	-4.521	2.563
	Sixth Grade	-.300	1.117	1.000	-3.842	3.242
	Seventh Grade	-.620	1.117	1.000	-4.162	2.922
	Eighth Grade	-.448	1.117	1.000	-3.990	3.094

Note. a. Adjustment for multiple comparisons: Bonferroni. Based on estimated marginal means.

I ran univariate tests because I wanted to determine if the data fit within the normal curve, which it did. Table 7 shows the results of the univariate test for females. The ANOVA indicated *F*(.168), and *p* = .984. Therefore, ANOVAs showed that grade is not a predictor for mathematical scores for females because there were no statistically significant differences.

Table 7

ANOVA Univariate Test Results for Female Students

	Sum of squares	*df*	Mean square	*F*	*Sig.*
Contrast	5.590	6	.932	.168	.984
Error	338.292	61	5.546		

Note. The F tests the effect of Grade. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Next, an ANOVA was conducted to analyze all of the males with grade as a fixed factor. Tables 8 to 11 show the results of the tests of between subject effects, estimates of grade levels, and univariate tests for male students. There were 10 representative grades for every grade except Grade 9. There were only 8 representative grades for Grade 9 because the students in Grade 9 were not administered the assessment in 2016.

Table 8 shows the results of the between-subjects effects tests. The dependent variable was percent. I ran a between-subjects effects test because it is an extension of the between groups *t* test to examine two or more groups simultaneously. The between-subjects factors were the grade levels. The between-subjects test indicated *F*(.056), *p* = .999. Therefore, grade level was not a determining factor for scores for males.

Table 8

ANOVA Test of Between-Subjects Effects for Male Students

Source	Type III sum of squares	*df*	Mean square	*F*	*Sig.*
Corrected model	5.598[a]	6	.933	.056	.999
Intercept	42941.286	1	42941.286	2571.414	.000
Grade	5.598	6	.933	.056	.999
Error	1018.668	61	16.699		
Total	44175.542	68			
Corrected total	1024.266	67			

Note. [a] R Squared = .005 (Adjusted R Squared = -.092).

Next, I ran an estimated marginal means test. The mean was 25.206 with a standard error of .497. The estimated marginal means test was run to reconfirm there were no outliers. There were no outliers. Then I ran the estimates for grade levels. Table 9 shows the results of the estimates for grade levels for male students. The estimates for grade levels test was run to see if there was a difference between the grade levels. The dependent variable was percent. There was not a difference between grade levels.

Table 9

ANOVA Estimates for Grade Levels for Male Students

			95% Confidence interval	
Grade	Mean	Std. error	Lower bound	Upper bound
Grade 3	25.194	1.292	22.610	27.778
Grade 4	24.945	1.292	22.361	27.529
Grade 5	24.751	1.292	22.167	27.335
Grade 6	25.430	1.292	22.846	28.014
Grade 7	25.111	1.292	22.527	27.695
Grade 8	25.282	1.292	22.698	27.866
Grade 9	25.731	1.445	22.842	28.620

Next, I ran the pairwise comparisons test. The dependent variable was labeled percent, and was the representative grade representing the percent of male students who scored at or above the national standard compared to the total number of students in the data set scoring at or above the national standard. The pairwise test was run to examine individual grades. There was no significant difference between individual grades. Table 10 shows the pairwise results for male students.

Table 10

ANOVA Pairwise Results for Male Students

(I) Grade	(J) Grade	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence interval Lower bound	Upper bound
Third Grade	Fourth Grade	.249	1.828	1.000	-5.545	6.043
	Fifth Grade	.443	1.828	1.000	-5.351	6.238
	Sixth Grade	-.236	1.828	1.000	-6.030	5.558
	Seventh Grade	.083	1.828	1.000	-5.711	5.877
	Eighth Grade	-.088	1.828	1.000	-5.882	5.706
	Ninth Grade	-.537	1.938	1.000	-6.683	5.608
Fourth Grade	Third Grade	-.249	1.828	1.000	-6.043	5.545
	Fifth Grade	.194	1.828	1.000	-5.600	5.989
	Sixth Grade	-.485	1.828	1.000	-6.279	5.309
	Seventh Grade	-.166	1.828	1.000	-5.960	5.628
	Eighth Grade	-.337	1.828	1.000	-6.131	5.457
	Ninth Grade	-.786	1.938	1.000	-6.932	5.359
Fifth Grade	Third Grade	-.443	1.828	1.000	-6.238	5.351
	Fourth Grade	-.194	1.828	1.000	-5.989	5.600
	Sixth Grade	-.679	1.828	1.000	-6.474	5.115
	Seventh Grade	-.360	1.828	1.000	-6.155	5.434
	Eighth Grade	-.531	1.828	1.000	-6.326	5.263
	Ninth Grade	-.981	1.938	1.000	-7.126	5.165
Sixth Grade	Third Grade	.236	1.828	1.000	-5.558	6.030
	Fourth Grade	.485	1.828	1.000	-5.309	6.279
	Fifth Grade	.679	1.828	1.000	-5.115	6.474
	Seventh Grade	.319	1.828	1.000	-5.475	6.113
	Eighth Grade	.148	1.828	1.000	-5.646	5.942
	Ninth Grade	-.301	1.938	1.000	-6.447	5.844
Seventh Grade	Third Grade	-.083	1.828	1.000	-5.877	5.711
	Fourth Grade	.166	1.828	1.000	-5.628	5.960
	Fifth Grade	.360	1.828	1.000	-5.434	6.155
	Sixth Grade	-.319	1.828	1.000	-6.113	5.475
	Eighth Grade	-.171	1.828	1.000	-5.965	5.623
	Ninth Grade	-.620	1.938	1.000	-6.766	5.525
Eighth Grade	Third Grade	.088	1.828	1.000	-5.706	5.882
	Fourth Grade	.337	1.828	1.000	-5.457	6.131
	Fifth Grade	.531	1.828	1.000	-5.263	6.326
	Sixth Grade	-.148	1.828	1.000	-5.942	5.646
	Seventh Grade	.171	1.828	1.000	-5.623	5.965
	Ninth Grade	-.449	1.938	1.000	-6.595	5.696
Ninth Grade	Third Grade	.537	1.938	1.000	-5.608	6.683
	Fourth Grade	.786	1.938	1.000	-5.359	6.932
	Fifth Grade	.981	1.938	1.000	-5.165	7.126
	Sixth Grade	.301	1.938	1.000	-5.844	6.447
	Seventh Grade	.620	1.938	1.000	-5.525	6.766
	Eighth Grade	.449	1.938	1.000	-5.696	6.595

Note. a. Adjustment for multiple comparisons: Bonferroni. Based on estimated marginal means.

Then I ran a univariate test because I wanted to determine if the data fit within the normal curve, which it did. The dependent variable was the percent of male students scoring at or above the standard. As indicated on page 118, there were 68 male representative grades used: two per year per grade level. The first representative grade for each year was the percentage of males scoring at the standard in relation to the total number of students scoring at or above the standard. The second representative grade for each year was the percentage of male students scoring above the standard in relation to the total number of students scoring at or above the standard. Table 11 shows the results from the univariate test for male students. The results were $df = 6$, $F(.056)$, and $p = .999$. Therefore, ANOVAs showed that grade is not a predictor for mathematical scores for males because there were no statistically significant differences.

Table 11

ANOVA Univariate Test Results for Male Students

	Sum of squares	*df*	Mean square	*F*	*Sig.*
Contrast	5.598	6	.933	.056	.999
Error	1018.668	61	16.699		

Note. The F tests the effect of Grade. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Finally, I conducted a logistical regression. Logistic regressions can provide an estimate of the probability of an outcome. The Logistical Regression was run to show the contributions of each independent variable to the model and its statistical significance. One independent variable entered was the standard (at or above the national average). The other independent variable was labeled percent, and was the representative grade representing the percentage of students who scored at or above the national standard compared to the total number of students in the data set scoring at or above the standard. During step 1 of the model summary for the logistical regression, the -2 Log likelihood= 186.809, the Cox & Snell R Square = .013, and the Nagelkerke R Square = .017. variables in the equation. The estimation was terminated at iteration number 3 because parameter estimates changed by less than .001.

Table 12 shows the results of the logistical regression variables in the equation.

Table 12

Logistical Regression Variables in the Equation

							95.0% Confidence interval for EXP (B)	
	B	*S.E.*	*Wald*	*df*	*Sig.*	*Exp(B)*	Lower bound	Upper bound
Percent	.134	.103	1.696	1	.193	1.143	.935	1.399
Standard	-.719	.650	1.224	1	.269	.487	.136	1.742
Constant	-2.271	1.829	1.542	1	.214	.103		

Note. Variable(s) entered on step 1: Percent, Standard.

For the independent variable labeled standard the results were *B*= .134, *df* = 1, *p* = .193, and *Exp (B)* = 1.143. For the independent variable labeled percent, which was the representative grade representing the percentage of students who scored at or above the national standard compared to the total number of students in the data set scoring at or above the standard, *B* = -.719, *df* = 1, *p* = .269, and *Exp (B)* = .487. Therefore, there was no statistical significance. No additional statistical tests were conducted because the null hypothesis was accepted for both research questions.

Summary

A correlational quantitative research design was used to show the relationship between gender and mathematics achievement. The statistical tests conducted were a *t* test, an ANOVA, and logistical regression. All of the statistical tests demonstrated there was no significant statistical evidence to support the research questions. Therefore, the null hypotheses were accepted for both research questions. The nature of the archival, de-identified data used in this study limited analysis. In addition, the data came from only one school district. Also, as military-connected students, the participants in this study may have been influenced by unique stressors such as frequent mobility and deployed parents as noted in Chapter 2. Therefore, results from this study should be interpreted with caution and should not be applied to other groups of military-connected students without careful consideration. In chapter 5 I concluded this dissertation by providing an interpretation of the results, limitations of the study, recommendations for future research, and implications for social change.

Chapter 5: Discussion, Conclusions, and Recommendations

Summary of the Study

Previous researchers examining large-scale national or international mathematics assessments have found gender inequity favoring males (Cheema & Galluzzo, 2013; Reilly et al., 2014; Stoet & Geary, 2013; Tyre, 2008). However, other researchers have found no significant differences in mathematics scores by gender (Kaleli-Yilmaz & Hanci, 2005; Lindberg et al., 2010). Some researchers found the largest gender gap in high achieving students (Hyde et al., 1990; Stoet & Geary, 2013), while others found grade level to be a factor, with the largest differences showing up at the high school level where males outperformed females (Lindberg et al., 2010). No studies were found during the literature review specifically examining elementary and post-elementary high achieving military-connected students in relation to gender and mathematics. Therefore, this study was conducted to add to the existing body of knowledge about high achieving military-connected students in relation to gender and mathematics because there is a gap in current research relating to these subjects.

The purpose of this quantitative study was to describe the relationship between gender and grade level to mathematics achievement on the TNTE for high achieving military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE using gender as a predictor and to determine if there was a statistically significant difference between the percentage of military-connected students scoring in the top two quartiles (at or above the national average) for NCE mathematics scores on the TNTE by gender and grade band. The student population was American high achieving students in Grades 3 through 9 with at least one parent affiliated with the U.S. military or working in a civilian occupation supporting the military between the years 2012 to 2016. All data came from a single participating school district that had a large population of military-connected students. The data were archival and de-identified by the participating school district's research center. The key findings from this study were that there were no statistical differences between the mean numbers of females and males by grade level or grade band, and, therefore, there was gender equity within the population studied.

Interpretation of Findings

In Chapter 2, I described the theoretical framework of Sax (2005, 2009, 2010, 2011). Sax conducted a meta-analysis of research on differences between male and female students. Sax identified gender differences in learning styles, vision, hearing, and behavior. Sax concluded anatomical, genetic, and functional differences between genders can impact learning and teachers should teach males and females differently in order to accommodate gender-specific needs. An important difference between the findings of this study and the theoretical framework of Sax (2005, 2009, 2010, 2011) is that analysis of the results of this study did not demonstrate a correlation between gender and mathematics for any of the grade levels or for either grade band.

In contrast to the results of this study, other researchers have found gender differences in mathematics achievement favoring male students (Cheema & Galluzzo, 2013; Kaleli-Yilmaz & Hanci, 2015), with the largest gap in high achievers at the 95th percentile and above (Stoet & Geary, 2013). Cheema and Galluzzo (2013) examined mathematics scores for American students in Grades 4, 8, and 12 and found a small but significant gender gap in mathematics performance with male students slightly outperforming female students on the PISA. Similarly, Stoet and Geary (2013) examined international mathematical data from 75 countries during the years 2000, 2003, 2006, and 2009 for approximately 1.5 million high school students and found that male students within the United States and across nations consistently scored higher than female students on the PISA. Furthermore, Stoet and Geary found that the largest gap was for high achieving students who scored at the 95th percentile and above. Similarly, while Lindberg et al. (2010) conducted a meta-analysis of over 1 million students and found negligible gender differences overall, Lindberg et al. did note a peak in high school with male students outperforming female students. Additionally, Reilly et al. (2014) conducted a meta-analysis of NAEP mathematical data for over a million students in Grades 4, 8, and 12 and found males moderately outperformed females in Grades 4 and 8, and there was a considerable gender gap in Grade 12 with males again outperforming females. Tyre (2008) examined NAEP scores for 17-year-old American students and discovered gender differences in mathematical achievement did exist with males outperforming females, but the difference had decreased from 8 points in 1973 to 3 points in 2004. Earlier, Hyde et al. (1990) conducted a meta-analysis of 100 studies over a 15-year time-period and reported the gender gap had shrunk since their study had begun but was still persistent. Hyde et al. noted that females outperformed males at the elementary and middle school levels, but males outperformed females at the high school and college level. Hyde et al. also found that the gender gap was largest for the highest achieving students. Similarly, Stoat and Geary (2013) and Reilly et al. (2014) also acknowledged that the persistent gender gap favoring males was small but still existent.

However, similar to this study, Kaleli-Yilmaz and Hanci (2015) also examined data from elementary students and post-elementary students. Kaleli-Yilmaz and Hanci found that gender was not a predictor of mathematical performance on the international TIMSS assessment for students in Grades 4, 8, and 11. Kaleli-Yilmaz and Hanci found that males performed higher than females in some countries and lower than females in others with students performing almost equally by gender in the United States. Kaleli-Yilmaz and Hanci determined that gender was a neutral element to mathematical scores. Similarly, Lindberg et al. (2010) conducted a meta-analysis of 20 years of literature from 242 studies and found no overall gender differences in elementary and middle school students in mathematical performance. Although the results of this study did not show a correlation between gender and mathematics, this study contributed to the knowledge base relating to high achieving military-connected students in relation to gender and mathematics.

Limitations of the Study

The data used in this study showed a balanced ratio in terms of female and male scores at all grade levels. There were several limitations to this study as noted in detail in Chapter 1 and are summarized as follows. The results of this study might not be generalizable to other populations because the student population from the participating school district selected for this study might not reflect the results of military-connected students from schools in other districts or from nonmilitary-connected students. This was a small study examining students from only one school district. The study was also limited by time and grade level examining only Grades 3 through 9 for a specific 5-year time period. In addition, only the high achieving students were examined.

A large limitation to this study was that I used de-identified secondary data from a highly mobile student population. I did not know the specific number of students or which students exited or entered individual schools or the participating school district as a whole during the years of the study, rather only year totals were available. Attrition could have affected internal validity. Similarly, the de-identified data did not include information about the testing environments, environmental factors that could have occurred during testing, or specific stressors that individual students faced at the time of testing, such as having a parent recently deployed to a war zone.

In addition, this study was limited because the de-identified data were not broken down into specific subgroups of students at each school such as those determined by poverty, race or ethnicity, parental education levels, parental marriage status, parental age, birth order, number of siblings, English Language Learner status, reading abilities, physical, mental, and emotional health, or special education needs. However, the results were consistent, showing gender equity for all grade levels and all years. Similarly, maturation was not a validity threat as there were no changes in test scores for different years or grades resulting in the knowledge that changes in the age of the participants was not a factor.

Finally, there are many other variables that could affect mathematics achievement specifically for military-connected students that were not specifically addressed, including but not limited to the following: parent deployment status, dual-military family status, branch of service, if one or more parents were currently or were preparing to be in a war zone, if a parent had military-related PTSD, number of stations a family had moved to, if a child was currently living overseas or in the United States, if a family was stationed near other family members, number of schools a child had attended, amount of time at the current station, if a child was living with someone other than the parents due to a dual-military deployment, whether the child lived on or off base, whether the parents were enlisted or officers, whether teachers and other staff were trained to specifically support the unique population of military-connected students within the schools, and whether students attended schools with a high population of military-connected students.

Recommendations for Future Research

The literature review for this study indicated that there are still few current studies that deal specifically with mathematics achievement and gender for military-connected students. More research should be conducted for educators to fully understand the role that gender and grade level play in mathematics achievement for military-connected students. The results of this study did not support the hypothesis of this research project. Although I anticipated that there could be gender differences, there were not. The results could be interpreted that the participating school system is doing something educationally correct because the school district was able to achieve gender equity in mathematics education for high achieving military-connected students in Grades 3 through 9. Therefore, future researchers could focus on examining possible reasons for the gender equity within the school system so that other school systems interested in increasing gender equity could benefit by replicating research-based methods of instruction or support within their own districts. Researchers might investigate whether the success of the schools in this study achieving gender equity in mathematics for high achieving military-connected students was due to teacher preparation, in-service trainings, or some other feature of the school district. A more in-depth examination of the participating school district in my study might include an examination of what training teachers are provided with not only to teach mathematics in such a way as to promote gender equity, but also what training teachers are provided to meet the unique stressors that military-children may face. In addition, future researchers might conduct an examination of the standards and curriculum that this district uses. Furthermore, researchers might examine the socioemotional and family supports that the participating school district has in place to meet unique needs of military-connected students. Results from such studies might reveal components that could be replicated in other districts with high populations of military-connected students.

Future researchers could also examine subgroups such as those determined by poverty, race or ethnicity, or English as a Second Language status within the highly mobile military-connected student population and compare those students to students in similar subgroups within the civilian population. Further study is also needed specifically examining the subgroup of gifted students. In addition, in this study, I only focused on students in the school district who scored at or above the standard in mathematics achievement. Further studies could address the students who scored below the standard in mathematics achievement.

A limitation to this study was that the data used were de-identified, and there was no way to track individual students to see if they continued to perform well in mathematics and if they chose STEM degrees at the university level and went on to join STEM professions. A long-term study following high achieving military-connected students through high school and university could provide more robust results.

I found the process of obtaining data on military-connected children to be challenging. It might be easier for researchers to obtain data on teachers of military-connected students because the teachers are adults and often civilians so there might be less restrictions to the process of studying them than there are on children. In Chapter 2 I noted a gap in the literature revealing a lack of longitudinal data on mathematics achievement of military-connected students in relation to gender and the lack of longitudinal data on military-connected children in public schools in general. A challenge I faced during this study was locating a school district with a high population of military-connected children because many public-school systems do not routinely identify or monitor this subset of children's enrollment, academic achievement (De Pedro et al., 2014; (Escueda et al., 2012). After I located a district that did identify and monitor military-connected students, I faced an additional challenge of obtaining permission to look at student scores because the students were military-connected. In order to overcome this challenge, I began by contacting an administrator at a school that I knew had a high population of military-connected students. I asked the administrator about the process of and difficulties that I might encounter obtaining data on military-connected students. Eventually I modified my original research questions and designs so that I could examine de-identified, archival data that had been pre-approved by the participating district's data research center because it sped up the process of obtaining data. I recommend that future researchers who are interested in doing research on military-connected student populations begin by locating a school district that has a high population of military-connected students and hold a discussion with administrators and the local military-school liaison to learn more about how to navigate the process of obtaining data on military-connected students or their teachers. Another avenue of obtaining data on military-connected students might be for researchers to contact the Department of Defense.

This study could be replicated using similar data from multiple school districts with large military-connected student populations and comparing elementary, middle school, and high school grade bands. Additional research involving more school districts and more grade levels may reveal new information in relation to grade level and gender equity in mathematics achievement for military-connected students. Additionally, it might be worth exploring the relationship of gender equity for military-connected American children nationally and living abroad to the following school-level factors: specific training for school staff on working with military families and connections between schools and military communities to support children and families with specific military related stressors. Other factors that could be studied are family income levels, parent education levels, and branch of service.

Finally, future research should investigate military-connected children who are dealing with stressors unique to military life such as children dealing with the following:

dual-military parents, parents deployed to war zones, parents deployed for extended periods of time, living with a limited-English-speaking parent while the other parent is deployed, moving multiple times resulting in a variety of schools, and living off base during an overseas assignment. All of these research ideas could provide researchers with a better understanding of the population of military-connected children.

Implications for Social Change

As I noted in Chapter 1, increasing the quantity of Americans choosing STEM professions is a national priority (Obama, 2005; United States Congress, 2015). Society benefits from achieving gender equity in the area of mathematics because gender equity could eliminate gender gaps at the graduate and post-graduate STEM degrees, which in turn lead to STEM professions. Similarly, eliminating gender gaps in STEM professions could increase America's socioeconomic competitiveness on the global scale. In order to eliminate gender gaps, studies need to be conducted to discover where they are and where they are not. The results of this study illuminated a school district that does not have gender inequity based on the results presented.

The findings of this study provide implications for positive social change in education as they contribute to the existing body of knowledge relating to gender, mathematics scores, and high achieving military-connected children in Grades 3 through 9. The results advance knowledge about overall student achievement in mathematics for American high achieving military-connected students. The results of this study also suggest the schools in the participating school district have found a way to teach mathematics in such a way that there was no gender inequity in the TNTE mathematics scores consistently over a 5-year period for students in Grades 3 through 9. Therefore, the results could be relevant to educational practitioners and researchers interested in gender-balanced mathematics education for military-connected students because they showed there is at least one school district that has gender equity.

Conclusion

The purpose of this study was to describe the relationships between gender and mathematics achievement on the TNTE for high achieving military-connected students in Grades 3 through 9 who were enrolled in schools in the participating school district. No statistically significant effects were found within the population studied.

These results may be helpful to school administrators, educators, and parents who wish to promote social change by contributing to a more balanced gender representation in higher level STEM enrollments and ultimately a more balanced gender representation in professional STEM related careers. These findings affirm military-connected students in the participating school district are receiving mathematical education resulting in no statistically significant achievement differences based on gender or gender and grade band, in Grades 3 through 9. Because gender differences do exist at higher levels of education and in STEM career fields, influential persons in these areas who are interested in gender equity may wish to examine the delivery modes of mathematics education in schools enrolling military-connected children. The results from this study provide implications for positive social change because they contribute to the body of knowledge about gender equity in mathematics for military-connected children. It is recommended that future studies consider other school districts with large military-connected populations to see if the aspect of being military-connected correlates with gender equity in mathematics on a larger scale or at the high school level.

References

Abell, C. S. (March 4, 2004). Prepared statement of the honorable Charles S. Abell principal deputy undersecretary of defense (personnel and readiness) before the Senate Armed Services Personnel Subcommittee. Retrieved from http://www.globalsecurity.org/military/library/congress/2004_hr/040304-Abell.pdf

Achiron, R., Lipitz, S., & Achiron, A. (2001). Sex-related differences in the development of the human fetal corpus callosum: In utero ultrasonographic study. *Prenatal Diagnosis, 21*(2), 116-120. doi:10.1002/1097-0223(200102)21:2<116: aid-pd19>3.3.co;2-d

Adabor, J. K. (2013). Harnessing formative and summative assessments to promote mathematical understanding and proficiency. *AURCO Journal*, *19*, 1956-66. Retrieved from http://aurco.net/journal_home

Alcock, L., Attridge, N., Kenny, S., & Inglis, M. (2014). Achievement and behavior in undergraduate mathematics: Personality is a better predictor than gender. *Research in Mathematics Education, 16*(1), 1-17. doi:10.1080/1479802.2013.874094

Alexander, G., & Hines, M. (2002). Sex differences in response to children's toys in nonhuman primates. *Evolution and Behavior, 23*(6) 467-479. doi:10.1016/s1090-5138(02)00107-1

Alon, S., & Gelbgiser, D. (2011). The female advantage in college academic achievements and horizontal sex segregation. *Social Science Research, 40*(1), 107-119. doi:10.1016/j.ssresearch.2010.06.007

Angrist, J., & Johnson, J. H. (2000). Effects of work-related absences on families: Evidence from the Gulf War. *Industrial and Labor Relations Review, 54*(1), 41. doi:10.2307/2696031

Arroyo, I., Burleson, W., Tai, M., Muldner, K., & Woolf, B. P. (2013). Gender differences in the use and benefit of advanced learning technologies for mathematics. *Journal of Educational Psychology*, *105*(4), 957-969. doi:10.1037/a0032748

Barker, L. H., & Berry, K. D. (2009). Developmental issues impacting military families with young children during single and multiple deployments. *Military Medicine, 174*(10), 1033-1040. doi:10.7205/MILMED-D-04-1108

Beardsley, S. (2015). DODEA math and reading scores among nation's highest. *Stars and Stripes.* Retrieved from http://www.stripes.com/news/DoDEA-math-and-reading-scores-among-nation-s-highest-1.375786

Bauerlein, M. (2008). *The dumbest generation: How the digital age stupefies young Americans and jeopardizes our future.* New York, NY: Jeremy P. Tarcher Inc.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences, 107*(5), 1860-1863. Retrieved from http://www.pnas.org/content/107/5/1860

Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science, 210*(4475), 1262-1264. doi:10.1126/science.7434028

Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science, 222*(4627), 1029-1031. doi:10.1126/science.6648516

Berenbaum, S., & Snyder, E. (1995). Early hormonal influences on childhood sex-typed activity and playmate preferences: Implications for the development of sexual orientation. *Developmental Psychology, 31*(1), 31-42. doi:10.1037//0012-1649.31.1.31

Birenbaum, M., Tatsuoka, C., & Xin, T. (2005). Large-scale diagnostic assessment: Comparison of eighth graders' mathematics performance in the United States, Singapore, and Israel. *Assessment in Education, 12*(2), 167-181. doi:10.1080/09695940500143852

Bradshaw, C. P., Sudhinaraset, M., Mmari, K., & Blum, R. (2010). School transitions among military adolescents: A qualitative study of stress and coping. *School Psychology Review, 39*(1), 84-105. Retrieved from http://naspjournals.org/?code=naps-site

Brown, R. S., & Coughlin, E. (2007). The predictive validity of selected benchmark assessments used in the Mid-Atlantic Region. Issues and Answers Report, REL 2007-No.017. Washington, D.C.: *U.S. Department of Education, Institute of Education Science, National Center for Education Evaluation and Regional Assistance, Regional Education Laboratory Mid-Atlantic.* Retrieved from https://ies.ed.gov/ncee/edlabs/regions/midatlantic/pdf/rel_2007017.pdf

Burrell, L. M., Adams, G. A., Durand, D. B., & Castro, C. A. (2006). The impact of military lifestyle demands on well-being, army, and family outcomes. *Armed*

Forces and Society, 33(1), 43–58. doi:10.1177/0002764206288804

Card, N. A., Bosch, L., Casper, D. M., Wiggs, C. B., Hawkins, S. A., Schlomer, G. L., & Borden, L. M. (2011). A meta-analytic review of internalizing, externalizing, and academic adjustment among children of deployed military service members. *Journal of Family Psychology, 25*(4), 508-520. doi:10.1037/a0024395

Casad, B. J., Hale, P., & Wachs, F. L. (2015). Parent-child math anxiety and math-gender stereotypes predict adolescents' math education outcomes. *Frontiers in Psychology, 6*, 1-12. doi:10.3389/fpsyg.2015.01597.

Casad, B. J., Hale, P., & Wachs, F. L. (2017). Stereotype threat among girls: Differences by gender identity and math education context. *Psychology of Women Quarterly, 41*(4), 513-529. doi:10.1177/0361684317711412

Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current Directions in Psychological Science*, *19*(5), 275-279. doi:10.1177/0963721410383241

Ceci, S. J., & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences, 108*(8), 3157-3162. doi:10.1073/pnas.1014871108

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135*(2), 218-261. doi:10.1037/a0014412

Chafin, A. F., Green, R. B., Raiford, S. A., E.-Ling, H., Nobles, K. D., & Truby, W. F. (2015). Using mathematics curriculum-based measurement to predict student performance on the third grade Georgia mathematics criterion referenced competency test. *National Teacher Education Journal*, *8*(2), 27-37. Retrieved from http://www.ntejournal.com/

Cheema, J. R., & Galluzzo, G. (2013). Analyzing the gender gap in math achievement: Evidence from a large-scale US sample. *Research in Education, 90*(1), 98-112. doi:10.7227/RIE.90.1.7

Cheryan, S., Ziegler, S. A., Montoya, A., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin, 143,* 1-35. doi:10.1037/bul0000052

Chu, M. W., Guo, Q., & Leighton, J. P. (2014). Students' interpersonal trust and attitudes towards standardized tests: Exploring affective variables related to student assessment. *Assessment in Education: Principles, Policy & Practice, 21*(2), 167-192. doi:10.1080/0969594X.2013/844094

College Board. (2007). 2007 *College-bound seniors: Total group profile report*. Retrieved from http://research.collegeboard.org/programs/sat/data/archived/cb-seniors-2007

College Board. (2015). *The 2015 SAT report on college and career readiness.* Retrieved from http://research.collegeboard.org/programs/sat/data/cb-seniors-2015

Committee on STEM Education: National Science and Technology Council. (2013). *A report from the committee on STEM education National Science and Technology Council. Executive Office of the President*. Retrieved from https://www.whitehouse.gov/sites/default/files/microsites/ostp/stem_stratplan_20

13.pdf

Cone-Wesson, B., & Ramirez, G. M. (1997). Hearing sensitivity in newborns estimated from ABRs to bone-conducted sounds. *Journal of the American Academy of Audiology, 8*, 299-307. Retrieved from http://www.audiology.org/

Connellan, J., Baron-Cohen, J., Wheelwright, S., Batki, A., & Ahluwalia, J. (2002). Sex differences in human neonatal social perception. *Infant Behavior & Development, 23*(1), 113-118. doi:10.1016/s0163-6383(00)00032-1

Contini, D., Di Tommaso, M. L., & Mendolia, S. (2017). The gender gap in mathematics achievement: Evidence from Italian data. *Economics of Education Review*, *58*(C), 32-42. doi:1016/j.econedurev.2017.03.001

Corso, J. (1959). Age and sex differences in pure-tone thresholds. *Journal of the Acoustical Society of America, 31(4)*, 489-507. doi:10.1121/1.1907742

Corso, J. (1963). Aging and auditory thresholds in men and women. *Archives of Environmental Heath, 6*(3), 350-356. doi:10.1080/00039896.10663405

Creech, S. K., & Hadley, W. (2014). The impact of military deployment and reintegration on parenting: A systematic review. *Professional Psychology: Research and Practice, 45*(6), 452-464. doi:10.1037/a0035055

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches. (3rd ed.).* Los Angeles, CA: SAGE Publications Inc.

CTB/McGraw-Hill. (2008). CTB McGraw-Hill's Terra Nova assessment recognized by U.S. department of education predictive validity research study: Study reviewed four benchmark assessment products used by states in the mid-Atlantic region –

TerraNova Ranked First. (Press Release). *PR Newswire.* Retrieved from http://www.prnewswire.com/news-releases/ctbmcgraw-hills-terranova-assessment-recognized-by-us-department-of-education-predictive-validity-research-study-57405852.html

CTB/McGraw-Hill. (2009). *Teacher's guide to TerraNova, third edition.* Monterey, CA: CTB/McGraw-Hill LLC.

Cutumisu, M., & Bulut, O. (2017). Problem-solving attitudes and gender as predictors of academic achievement in mathematics and science for Canadian and Finnish students in the PISA 2012 Assessment. In J. Johnson (Ed.), *Proceedings of EdMedia 2017*, 728-738. Washington, DC: Association for the Advancement of Computing in Education (AACE). Retrieved from https://www.learntechlib.org/primary/p/178382

D'Alessandro, L. M., & Norwich, K. H. (2009). Loudness adaptation measured by the simultaneous dichotic loudness balance technique differs between genders. *Hearing Research, 247*(2), 122-127. doi:10.1016/j.heares.2008.10.009

De Pedro, K. T., Astor, R. A., Benbenishty, R., Estrada, J., Dejoie Smith, G. R., & Esqueda, M. C. (2011). The children of military service members: Challenges, supports, and future educational research. *Review of Educational Research, 81*(4), 566-618. doi:10.3102/0034654311423537

De Pedro, K. T., Esqueda, M. C., Cederbaum, J. A., & Astor, R. A. (2014). District, school, and community stakeholder perspectives of the experiences of military-connected students. *Teachers College Record*, *116*(5), 1-32. Retrieved from http://www.tcrecord.org/library/Abstract.asp?ContentId=17438

Department of Defense Education Activity. (2010). *About DoDEA*. Retrieved from http://www.DoDEA.edu/home/about.cfm

Department of Defense Education Activity. (2011a). *Deployment and transition support*. Retrieved from http://www.DoDEA.edu/back_to_school/2010_11.cfm?cId=Deployment

Department of Defense Education Activity. (2011b). *Military K–12 partners: A DoDEA partnership program.* Retrieved from http://www.militaryk12partners.DoDEA.edu/index.cfm

Department of Defense Education Activity. (2014a). *Community strategic plan volume 1: School years 2013/14-2017/18*. Retrieved from http://www.DoDEA.edu/CSP/upload/CSP_130703.pdf

Department of Defense Education Activity. (2014b). *Community strategic plan volume 2: School years 2013/14-2017/18*. Retrieved from http://www.DoDEA.edu/CSP/upload/DoDEA_CSP_Vol2_CMYK_printready.pdf

Department of Defense Education Activity. (2015). *About DoDEA: DoDEA schools worldwide!* Retrieved from http://www.DoDEA.edu/aboutDoDEA/today.cfm

Diprete, T. A., & Buchmann, C. (2013). *The rise of women: The growing gender gap in education and what it means for American schools*. New York, NY: Russell Sage Foundation.

Doig, B. (2006). Large-scale mathematics assessment: Looking globally to act locally. *Assessment in Education: Principles, Policy, & Practice, 13*(3), 265-288.

doi:10.1080/09695940601035403

Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78-89. doi:10.1080/00461520902832368

Elliot, C. (1971). Noise tolerance and extraversion in children. *British Journal of Psychology, 62*(3), 375-380. doi:10.1111/j.2044-8295.1971.tb02048.x

Engel, R. C., Gallagher, L. B., & Lyle, D. S. (2010). Military deployments and children's academic achievement: Evidence from Department of Defense Education Activity schools. *Economics of Education Review, 29*(1), 75–82. doi:10.1016/j.econedurev.2008.12.003

Esqueda, M. C., Astor, R. A., & De Pedro, K. M. (2012). A call to duty: Educational policy and school reform addressing the needs of children from military families. *Educational Researcher, 31*(2), 65-70. doi:10.3102/0013189X11432139

Force, A. T. (2007). *Report of the APA task force on the sexualization of girls*. Washington, D.C.: American Psychological Association. Retrieved from http://www.apa.org/pi/women/programs/girls/report-full.pdf

Frankfort-Nachmias, C. & Nachmias, D. (2008). *Research methods in the social sciences*. New York, NY: Worth Publishers.

Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual differences, 47*, 182-193. doi:10.1016/j.lindif.2016.01.002

Garner, J. K., Arnold, P.L., & Nunnery, J. (2014). Schoolwide impact of military-connected student enrollment. Educators' perceptions. *Children & Schools, 36*(1). 31-39. doi:10.1093/cs/cdt026

Gaspard, H., Dicke, A. L., Flunger, B., Schreier, B. Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*(3), 663-677. doi:10.1037/edu0000003

Gibbs, D. A., Martin, S. L., Kupper, L. L., & Johnson, R. E. (2007). Child maltreatment in enlisted soldiers' families during combat-related deployment. *JAMA, 298*(5), 528-535. doi:10.1001/jama.298.5.528

Grön, G., Wunderlich, A. P., Spitzer, M., Tomczak, R., & Riepe, M. W. (2000). Brain activation during human navigation: Gender different neural networks as substrate of performance. *Nature Neuroscience, 3*(4), 404-408. doi:10.1038/73980

Grunspan, D. Z., Eddy, S. L., Brownell, S. E., Wiggins, B. L., Crowe, A. J., & Goodreau, S. M. (2016). Males under-estimate academic performance of their female peers in undergraduate biology classrooms. *PLoS ONE, 11*(2), e0148405. doi:10.1371/journal.pone.0148405

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S., & Yeung, A. S. (2015). Expectancy-value in mathematics, gender and socioeconomic background as predictors of achievement and aspirations: A multi-cohort study. *Learning and Individual Differences, 37*, 161-168. doi:10.1016/j.lindif.2015.01.008

Gur, R. C., Turetsky, B. I., Matsui, M., Yan, M., Bilker, W. Hughett, P., & Gur, R. E. (1999). Sex differences in brain gray and white brain matter in healthy young adults: Correlations with cognitive performance. *The Journal of Neuroscience, 19*(10), 4065-4072. doi:10.1097/00019442-200201000-00009

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*(1), 1-51. doi:10.1111/j.1529-1006.2007.00032x

Hanlon, H., Thatcher, R., & Cline, M. (1999). Gender differences in the development of EEG coherence in normal children. *Developmental Neuropsychology, 16*(3), 479-506. doi:10.1207/s15326942dn1603_27

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*(1), 33-46. doi:10.2307/749455

Hill, F., Mammarella, I.C., Devine, A., Caviola, S., Passolunghi, M.C., & Szűcs, D. (2016). Maths anxiety in primary and secondary school students: Gender differences, developmental changes and anxiety specificity. *Learning and Individual Differences*, *48*, 45-53. doi:10.1016/j.lindif.2016.02.006

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*(6), 581-592. doi:10.1037/0003-066x.60.6.581

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139-155. doi:10.1037/0033-2909.107.2.139

IBM Corp. (2016). *IBM SPSS statistics for Windows, Version 24.0*. Armonk, NY: IBM Corp.

Johnson, H. J., Barnard-Brak, L., Saxon, T. F., & Johnson, M. K. (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *The Journal of Experimental Education, 80*(2), 137-149. doi:10.1080/00220973.2011.567312

Jones, J. (2010). Closing the gender gap. *Civil Engineering Magazine Archive, 80*(7), 60-63. doi:10.1061/ciegag.0000304

Kaleli-Yilmaz, G., & Hanci, A. (2015). Examination of the 8th grade students' TIMSS mathematics success in terms of different variables. *International Journal of Mathematical Education in Science and Technology, 47*(5), 674-695. doi:10.1080/0020739X.2015.1102977

Kao, C. (2015). Mathematically gifted adolescent females' mixed sentiment toward gender stereotypes. *Social Psychology of Education, 18*(1), 17-35. doi:10.1007/s11218-014-9278-2

Kaplan, E., & Benardete, E. (2001). The dynamics of primate retinal ganglion cells. *Progress in Brain Research, 134*, 17-34. doi:10.1016/s0079-6123(01)34003-7

Krishnamurthi, A., Bevan, B., Rinehart, J., & Coulon, V. R. (2013). What afterschool STEM does best: How stakeholders describe youth learning outcomes. *Afterschool Matters, (18),* 42-49. Retrieved from https://www.niost.org/Afterschool-Matters/afterschool-matters-journal

Laursen. S. L., Thiry, H., Archie, T., & Crane, R. (2013). Variations on a theme: Characteristics of out-of-school time science programs. *Afterschool Matters, (17),* 36-49. Retrieved from https://www.niost.org/Afterschool-Matters/afterschool-matters-journal

Lemmon, K., & Stafford, E. (2014). Advocating for America's military children: Considering the impact of parental combat deployment to Iraq and Afghanistan. *Family Court Review, 52*(3), 343-354. doi:10.1111/fcre.12096

Lester, P., Aralis, H., Sinclair, M., Kiff, C., Lee, K., Mustillo, S., & Wadsworth, S. M. (2016). The impact of deployment on parental, family, and child adjustment in military families. *Child Psychiatry & Human Development, 47*(6), 938-949. doi:10.1007/s10578-016-0624-9

Lester. P., & Flake, E. Lieutenant Colonel U.S. Air Force. (2013). How wartime military service affects children and families. *The Future of Children, 23*(2), 121- 141. doi:10.1353/foc.2013.0015

Lester, P., Peterson, K., Reeves, J., Knauss, L., Glover, D., Mogil, C., Beardslee, W. (2010). The long war and parental combat deployment: Effects on military children and at-home spouses. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(4), 310-320. doi:10.1016/j.jaac.2010.01.003

Lever, J. (1976). Sex differences in the games children play. *Social Problems, 23*(4), 478-487. doi:10.1525/sp.1976.23.4.03a00100

Lever, J. (1978). Sex differences in the complexity of children's games. *American Sociological Review, 43*(4), 471-483. doi:10.2307/2094773

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*(6), 1123-1135. doi:10.1037/a0021276

Lucier-Greer, M., Arnold, A. L., Mancini, J. A., Ford, J. L., & Bryant, C. M. (2015) Influences of cumulative risk and protective factors on the adjustment of adolescents in military families. *Family Relations, 64*(3), 363-377. doi:10.1111/fare.12123

Lyle, D. S. (2006). Using military deployments and job assessments to estimate the effects of parental absences and household relocations on children's academic achievement. *Journal of Labor Economics, 24*(2), 318-350. doi:10.1086/499975

Maestripieri, D., & Pelka, S. (2002). Sex differences in interest in infants across the lifespan: A biological adaptation for parenting? *Human Nature, 13*(3), 327-344. doi:10.1007/s12110-002-1018-1

Mangels, J. A., Good, C., Whiteman, R. C., Maniscalco, B., & Dweck, C. S. (2012). Emotion blocks the path to learning under stereotype threat. *Social Cognitive and Affective Neuroscience, 7*(2), 230-241. doi:10.1093/scan/nsq100

Mann, A., & DiPrete, T. A. (2013). Trends in gender segregation in the choice of science and engineering majors. *Social Science Research, 42*(6), 1519-1541. doi:10.2139/ssm.1938328

Martin, C. (2007). *Perfect girls, starving daughters: The frightening new normalcy of hating your body.* New York, NY: Free Press.

Martinez, S., Mizala, A., Martinez, F, … & ... Martinez, S. (2016). Pre-service elementary school teachers' expectations about student performance: How their beliefs are affected by their mathematics anxiety and student's gender. *Teaching and Teacher Education*, *50,* 70-78. doi:10.1016/j.tate.2015.04.006

Marx, D. M., & Stapel, D. A. (2006). Understanding stereotype lift on the role of the social self. *Social Cognition 23*(6), 776-791. doi:10.1521/soco.2006.24.6.776

McGlone, J. (1980). Sex differences in human brain asymmetry: A critical survey. *Behavioral and Brain Sciences, 3*(2), 215-263. Retrieved from doi:10.1017/s0140525x00004398

McGuinness, D. (1974). Equating individual differences for auditory input. *Psychophysiology, 11*(2), 113-120. doi:10.1111/j.1469-8986.1974.tb00831.x

McLeod, S. (2008). *Correlation.* Retrieved from http://www.simplypsychology.org/correlation.html

Meissirel, C., Wikler, K., Chalupa, L., & Rakic, P. (1997). Early divergence of magnocellular and parvocellular functional subsystems in the embryonic primate visual system. *Proceedings of the National Academy of Sciences, 94,* (11), 5900-5905. doi:10.1073/pnas.94.11.5900

Mendick, H. (2005). Only connect: Troubling oppositions in gender and mathematics. *International Journal of Inclusive Education, 9*(2), 161-180. doi:10.1080/1360311042000339383

Military.com. (n.d.) *U.S. armed forces overview*. Retrieved from http://www.military.com/join-armed-forces/us-military-overview.html

Military.com. (2015). *Deployment: An overview*. Retrieved from http://www.military.com/deployment/deployment-overview.html

Military Onesource. (2015). 2015 demographics: Profile of the military community. Retrieved from http://download.militaryonesource.mil/12038/MOS/Reports/2015-

Demographics-Report.pdf

Moeller, J. D., Culler, E. D., Hamilton, M. D., Aronson, K. R., & Perkins, D. F. (2015). The effects of military-connected parental absence on the behavioural and academic functioning of children: A literature review. *Journal of Children's Services, 10*(3), 291-306. doi:10.1108/JCS-05-2015-0017

Morgan, S. L., Gelbgiser, D., & Weeden, K. A. (2013). Feeding the pipeline: Gender, occupational plans, and college major selection. *Social Science Research, 42*(4), 989-1005. doi:10.1016/j.ssresearch.2013.03.008

Mosatche, H. S., Matloff-Nieves, S., Kekelis, L., & Lawner, E. K. (2013). Effective STEM programs for adolescent girls: Three approaches and many lessons learned. *Afterschool Matters, (17)*, 17-25. Retrieved from https://www.niost.org/Afterschool-Matters/afterschool-matters-journal

National Center for Education Statistics. (2015). *2015 mathematics grades 4 and 8 assessment report cards: Summary data tables for national and state average scores and achievement level results*. Retrieved from http://www.nationsreportcard.gov/reading_math_2015/files/2015_Results_Appendix_Math.pdf

National Science Foundation. (2011). *Women, minorities, and persons with disabilities in science and engineering*. Arlington, VA: National Science Foundation. Retrieved from https://www.nsf.gov/statistics/2017/nsf17310/static/downloads/nsf17310-digest.pdf

National Science Foundation. (2017). *Women, minorities, and persons with disabilities in* *science and engineering: 2017.* (Special Report NSF 17-310). Arlington, VA: National Center for Science and Engineering Statistics.

Negrusa, S., Negrusa, B., & Hosek, J. (2014). Gone to war: Have deployments increased divorces? *Journal of Population Economics, 27*(2), 473-496. doi:10.1007/s00148-013-0485-5

Obama, B. (2005). *Literacy and education in a 21st-century economy.* Retrieved from http://obamaspeeches.com/024-Literacy-and-Education-in-a-21st-Century-Economy-Obama-Speech.htm

O'Gara, F., & Kanellis, E. (2015). *DODEA ranks with top performing states in the nation on the 2015 NAEP*. Retrieved from http://www.DoDEA.edu/newsroom/pressreleases/2015-NAEP.cfm

Ogueta, S. B., Schwartz, S. D., Yamashita, C. K., & Farber, D. B. (1999). Estrogen receptor in the human eye: Influence of gender and age on gene expression. *Investigative Ophthalmology & Visual Science 40*(9), 1906-1911.Retrieved from http://iovs.arvojournals.org/article.aspx?articleid=2122996

Pahlke, E., Hyde, S., & Allison, C. M. (2014). The effects of single-sex compared with coeducational schooling on students' performance and attitudes: A meta-analysis. *American Psychological Association Psychological Bulletin 140*(4), 1042-1072. doi:10.1037/a0035740

Picho, K., Rodriguez, A., & Finnie, L. (2013). Exploring the moderating role of context on the mathematics performance of females under stereotype threat: A meta-analysis. *The Journal of Social Psychology, 153*(3), 299-333.

doi:10.1080/00224545.2012.737380

Plante, I., Théorêt, M., & Favreau, O. E. (2009). Student gender stereotypes: Contrasting the perceived maleness and femaleness of mathematics and language. *Educational Psychology, 29*(4), 285-405. doi:10.1080/01443410902971500

Pomerantz, E., Altermatt, E., & Saxon, J. (2002). Making the grade but feeling distressed: Gender differences in academic performance and internal distress. *Journal of Educational Psychology, 94*(2), 396-404. doi:10.1037/0022-0663.94.2.396

Pomerantz, E. & Saxon, J. (2001). Conceptions of ability as stable and self-evaluative processes: A longitudinal examination. *Child Development, 72(1),* 152-173. doi:10.1111/1467-8624.00271

Reilly, D., Neumann, D. L., & Andrews, G. (2014). Sex differences in mathematics and science achievement: A meta-analysis of national assessment of educational progress assessments. *Journal of Educational Psychology, 107*(3), 645-662. doi:10.1037/edu0000012

Richardson, E. E., Mallette, J., O'Neal, C., & Mancini, J. (2016). Do youth development programs matter? An examination of transitions and well-being among military youth. *Journal of Child & Family Studies, 25*(6), 1765-1776. doi:10.1007/s10826-016-0361-5

Rogers, D. S., Harkrider, A. W., Burchfield, S. B., & Nabelek, A. K. (2003). The influence of listener's gender of the acceptance of background noise. *Journal of the American Academy of Audiology, 14*, 372-382. Retrieved from http://www.audiology.org/publications/about-journal-american-academy-audiology

Sagi, E., D'Alessandro, L. M., & Norwich, K. H. (2007). Identification variability as a measure of loudness: An application to gender differences. *Canadian Journal of Experimental Psychology, 61*(1), 64-70. doi:10.1037/cjep2007007

Sathyanarayana, S., Karr, C.J., Lozano, P., Brown, E., Calafat, A. M., Liu, F., & Swan, S. H. (2008). Baby care products: Possible sources of infant phthalate exposure. *Pediatrics, 121*(2), e260-268. doi:10.1542/peds.2006-3766

Sax, L. (2005). *Why gender matters: What parents and teachers need to know about the emerging science of sex differences.* New York, NY: Doubleday.

Sax, L. (2009). *Boys adrift: The five factors driving the growing epidemic of unmotivated boys and underachieving young men.* New York, NY: Basic Books.

Sax, L. (2010). Sex differences in hearing: Implications for best practice in the classroom. *Advances in Gender and Education, 2*, 13-21. Retrieved from http://www.leonardsax.com/sex-differences-in-hearing

Sax, L. (2011). *Girls on the edge: The four factors driving the new crisis for girls: Sexual identity, the cyberbubble, obsessions, environmental toxins.* New York, NY: Basic Books.

Sax, L. (2016). *Why gender matters: Chapter updates*. [Web log post]. Retrieved from http://www.leonardsax.com/books/why-gender-matters/

Sax, L. J., Kanny, M. A., Riggers-Piehl, T. A., Whang, H., & Paulson, L. N. (2015). "But I'm not good at math:" The changing salience of mathematical self-concept in shaping women's and men's STEM aspirations. *Research in Higher Education,* *56*(8), 813-842. Retrieved from https://doi.org/10.1007/s11162-015-9375-x

Schoenfeld, A. H. (2015). Summative and Formative Assessments in Mathematics Supporting the Goals of the Common Core Standards. *Theory Into Practice*, *54*(3), 183-194. doi:10.1080/00405841.2015.1044346

Schommer-Aikens, M., Unruh, S., & Morphew, J. (2015). Epistemological belief congruency in mathematics between vocational technology students and their instructors. *Journal of Education and Training Studies, 3*(4), 137-145. doi:10.11114/jets.v3i4.859

Shuttleworth, M. (2009). Construct validity. *Explorable.com.* Retrieved from https://explorable.com/construct-validity

Silverthorne, S. (2009). Understanding users of social networks. *Harvard Business School, 14*. Retrieved from http://hbswk.hbs.edu/item/understanding-users-of-social-networks

Soper, D. (2017). Calculator: A-priori sample size for student t-tests. Retrieved from http://www.danielsoper.com/statcalc/calculator.aspx?id=47

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28. Retrieved from https://www.sciencedirect.com/journal/journal-of-experimental-social-psychology

Spitzer, B., & Aronson, J. (2015). Minding and mending the gap: Social psychological interventions to reduced educational disparities. *British Journal of Educational Psychology, 85*(1), 1-18. doi:10.1111/bjep.12067

Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811. doi:10.1037//0022-3514.69.5.797

Steingraber, S. (2007). The falling age of puberty in U.S. girls: What we know, what we need to know. *San Francisco Breast Cancer Fund.* Retrieved from http://www.breastcancerfund.org/assets/pdfs/publications/falling-age-of-puberty.pdf 57.

Stoet, G., Bailey, D. H., Moore, A. M., & Geary, D. C. (2016). Countries with higher levels of gender equality show larger national sex differences in mathematics anxiety and relatively lower parental mathematics valuation for girls. *PloS one, 11*(4), e01538. doi:10.137`/journal.pone.0153857

Stoet, G., & Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology, 16*(1), 93-102. doi:10.1037/a0026617

Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within-and across-nation assessment of 10 years of PISA data. *PLoS One*, *8*(3), e57988. Retrieved from doi:10.1371/journal.pone0057988

Stoet, G., & Geary, D. C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological Science, 29*(4), 581-593. doi:10.1177/0956797617741719

Tatar, E., Zengin, Y., & Kağızmanlı, T. (2015). What is the relationship between technology and mathematics teaching anxiety. *Journal of Educational Technology & Society*, *18*(1), 67-76. Retrieved from http://www.ifets.info/

Tichenor, M., Welsh, A., Corcoran, C., Piechura, K., & Heins, E. (2016). Elementary girls' attitudes toward mathematics in mixed-gender and single-gender classrooms. *Education*, *137*(1), 93-100. Retrieved from http://www.projectinnovation.com/

Tine, M., & Gotlieb, R. (2013). Gender-, Race-, and Income-Based Stereotype Threat: The Effects of Multiple Stigmatized Aspects of Identity on Math Performance and Working Memory Function. *Social Psychology of Education: An International Journal*, *16*(3), 353-376. doi:10.1007/s11218-013-9224-8

Tolley, K. (2003). *The science education of American girls: A historical perspective.* London, England: Routledge.

Tyler-Wood, T., Ellison, A., Lim, O., & Periathiruvadi, S. (2012). Bringing up girls in science (BUGS): The effectiveness of an afterschool environmental science program for increasing female students' interest in science careers. *Journal of Science Education and Technology, 21*(1), 46-55. doi:10.1007/s10956-011-9279-2

Tyre, P. (2008). *The trouble with boys: A surprising report card on our sons, their problems at school, and what parents and educators must do.* New York, NY: Crown.

UNESCO Institute for Statistics. (2015). *Education: Percentage of female teachers by teaching level of education.* Retrieved from http://data.uis.unesco.org/index.aspx?queryid=178

United States Census Bureau. (2012). *2010 U.S. Census Population Data.* Retrieved from https://www.census.gov/2010census/

United States Congress. (2015). *Reintroduction of the STEM education act. H.R. 1020. 114th Congress 1st session.* Retrieved from https://www.congress.gov/114/bills/hr1020/BILLS-114hr1020ih.pdf

Valeski, T., & Stipek, D. (2001). Young children's feelings about school. *Child Development, 72*(4), 1198-1213. doi:10.1111/1467-8624.00342

Wagemaker, H. (2002). *TIMSS in context: Assessment, monitoring, and moving targets,* in Robitaille, D.F., & Beaton, A.E. (Eds) *Secondary Analysis of TIMSS data.* Boston, MA: Springer Books. Retrieved from https://www.springer.com/la/book/9781402008597

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study. *Journal of Educational Psychology, 102*(4), 860-871. doi:10.1037/a0019454

Walton, G., & Cohen, G. (2003). Stereotype lift. *Journal of Experimental Psychology 39*(5),456-467. doi:10.1016/s0022-1031(03)00019-2

Wang, M., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review, 29*(1), 119-140. doi:10.1007/s10648-015-9355-x

Wang, M., Eccles, J. S., & Kenny, S. (2013). Not lack of ability but more choice: Individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychological Science, 24*(5), 770-775. doi:10.1177/0956797612458937

Waters, J. (n.d.). Correlational research guidelines: Conducting correlational research. *Capilano University*. Retrieved from https://www.capilanou.ca/psychology/student-resources/research-guidelines/Correlational-Research-Guidelines/

Watt, H. G., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., & Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: A comparison of samples from Australia, Canada, and the United States. *Developmental Psychology, 48*(6), 1594-1611. doi:10.1037/a0027838

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J. (2003). *Looking inside the classroom: A study of K-12 mathematics and science education in the United States.* Chapel Hill, NC: Horizon Research, Inc.

Wichstrom, L. (1999). The emergence of gender differences in depressed mood during adolescence: The role of intensified gender socialization. *Developmental Psychology, 35*(1), 232-245. Retrieved from https://doi.org/10.1037//0012-1649.35.1.232

Wickham, L. A., Gao, J., Toda, I., Rocha, E. M., Ono, M., & Sullivan, D. A. (2000). Identification of androgen, estrogen and progesterone receptor mRNAs in the eye.

Acta Ophthalmologica Scandinavica, 78(2), 146-153. Retrieved from https://onlinelibrary.wiley.com/journal/17553768

Wilkerson, S. B., & Haden, C. M. (2014). Effective practices of evaluating STEM out-of-school-time programs. *Afterschool Matters, (19),* 10-19. Retrieved from https://www.niost.org/Afterschool-Matters/afterschool-matters-journal

Yenilmez, K., & Turgut, M. (2016). Relationship between prospective middle school mathematics' teachers logical and reflective thinking skills. *Journal of Educational and Instructional Studies in the World, 6*(4), 15-20. Retrieved from http://www.wjeis.org/

Yin, R. K. (2009). *Case study research: Design and methods. (4th Ed.).* Thousand Oaks, CA: SAGE Publications.

Yin, R. K. (2012). *Applications of case study research.* (3rd Ed.). Thousand Oaks, CA: SAGE Publications.

Zorrilla-Silvestre, L., Presentación-Herrero, M. J., & Gil-Gómez, J. (2016). The relationship between neuropsychological and ecological measurements of executive functioning in childhood and the prediction of mathematics performance. A pilot study. *Electronic Journal of Research in Educational Psychology*, *14*(2), 333-351. doi:10.14204/ejrep.39.15080

Appendix A: Letter of Cooperation

The participating school district’s Research Center provided me with a letter of cooperation but requested that all identifying information be removed before publication of this dissertation. The original letter of cooperation is on file with Walden University.